National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Services des thèses canadiennes

Ottawa, Canada
K1A 0N4

CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

Canada

NL-339 (r.86/06)

THE UNIVERSITY OF ALBERTA

NUMERICAL MODELING OF CMOS MAGNETIC FIELD SENSORS BY FINITE ELEMENT METHOD

by

YOUNGSONG MUN

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

DEPARTMENT OF ELECTRICAL ENGINEERING

EDMONTON, ALBERTA

SPRING,1986

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-30161-9

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR

YOUNGSONG MUN

TITLE OF THESIS NUMERICAL MODELING OF CMOS MAGNETIC FIELD SENSORS BY FINITE ELEMENT METHOD

DEGREE FOR WHICH THESIS WAS PRESENTED MASTER OF SCIENCE

YEAR THIS DEGREE GRANTED SPRING,1986

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) Mun Youngsong

PERMANENT ADDRESS:

706-39, SINRIM3-DONG, GWANAK-GU,

SEOUL, KOREA

DATED 1, MAY, 1986

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled NUMERICAL MODELING OF CMOS MAGNETIC FIELD SENSORS BY FINITE ELEMENT METHOD submitted by YOUNGSONG MUN in partial fulfilment of the requirements for the degree of MASTER OF SCIENCE.

.....

Supervisor

.....

.....

Date 30 APR 86

ABSTRACT

Numerical solutions for magnetic field sensitive semiconductor devices such as VHD (Vertical Hall effect Device) and MAGFET (MAGnetic-field sensitive Field Effect Transistor) are developed. Poisson type partial differential equations having mixed boundary conditions for electrostatic potential are solved using the finite element method.

To support the finite element method efficiently, an adaptive grid generation algorithm is developed. The optimum sensor probe position for the VHD is found and the corresponding sensitivity is calculated. Double drain and Triple drain n-channel MAGFET operating in the linear region are simulated and their sensitivity are compared.

ACKNOWLEDGEMENTS

I wish to appreciate most sincerely to my supervisor Professor H. P. Baltes for his patient supervision and invaluable advice.

I would also like to thank Professors W. Allegretto, W. B. Joerg, and Dr. A. M. J. Huiser for their helpful assistance and discussions. I am indebted to a number of colleagues, especially A. Nathan for the simulating discussions, D. Briglio and K. Chau for proofreading this thesis.

Finally, special thanks are due to Mrs. June Swanson, who typed all the equations.

Table of Contents

Chapter		Page
1.	INTRODUCTION	1
2.	MAGNETIC FIELD EFFECTS IN SEMICONDUCTOR	6
	2.1 Hall Effect	6
	2.2 Carrier Deflection	7
	2.3 Magnetoconcentration	7
	2.4 Hall Angle	8
	2.5 Ideal Hall Voltage	8
3.	BASIC SEMICONDUCTOR EQUATIONS AND PREVIOUS MODELING WORKS	10
	3.1 Current Density Equations	10
	3.1.1 Current density equations without magnetic field	10
	3.1.2 Current density equations with magnetic field	11
	3.1.3 Current density components with magnetic field	13
	3.2 Current Continuity Equations	14
	3.3 Poisson's Equation	15
	3.4 Modeling by Lumped Circuit Cell Method	15
	3.5 Modeling by the Finite Difference Method	16
	3.6 Analytical Modeling	18
4.	GRID GENERATION	19
	4.1 Basic Concept	19
	4.1.1 Comparison of the finite element method and the finite difference method	19
	4.1.2 The reason for grid refinement	19
	4.2 Grid Generation Procedure	21
	4.2.1 Initial grid generation	21
	4.2.2 Numbering scheme for an initial grid	22
	4.2.3 Grid refinement	23

4.2.3.1 Horizontal refinement pass 24
4.2.3.2 Vertical refinement pass 25
4.2.4 Criteria to exit from grid generation 25
4.3 Results and discussions 26
5. FINITE ELEMENT METHOD 28
5.1 Introduction 28
5.2 Finite Element Method for Poisson's Equation 30
5.2.1 Interpolation functions 32
5.2.2 Assemblage of the local matrices 39
6. ANALYSIS OF THE VERTICAL HALL EFFECT DEVICES 41
6.1 Introduction. 41
6.2 Equations 41
6.3 Solution Procedure 42
6.4 Results and Discussion 44
7. ANALYSIS OF THE MAGFET 46
7.1 Assumptions 46
7.2 Basic Equations 47
7.3 Boundary Conditions 51
7.4 Numerical Method 51
7.5 Solution Procedure 52
7.6 Results and Discussion 52
8. CONCLUSION AND OUTLOOK 54
9. REFERENCES 95
10. APPENDIX 101

LIST OF TABLES

Table	Description	Page
6.1	Position of the electrodes	56
6.2	Voltages and currents of the electrodes at position 6 (μ*B = 0.2)	56

LIST OF FIGURES

Figure	Page
Fig. 2.1 (a) The Hall effect in a p-type semiconductor	57
(b) The Hall effect in a n-type semiconductor	
Fig. 2.2 Equipotential and current lines (L = 4W)	58
Fig. 2.3 Equipotential and current lines (W = 4L)	58
Fig. 2.4 Equipotential and current lines of intrinsic square device	58
Fig. 4.1 The numbering scheme for the initial grid	59
Fig. 4.2 Refinable element shapes for horizontal refinement pass	60-63
Fig. 4.3 A grid for a rectangular region	
(a) An initial grid	64
(b) A final grid with 400 nodes	65
Fig. 4.4 First example of a grid for a region having flat bottom	66
(a) An initial grid	
(b) A final grid with 200 nodes	
Fig. 4.5 Second example of a grid for a region having flat bottom	67
(a) An initial grid	
(b) A final grid with 230 nodes	
Fig. 4.6 Third example of a grid for a region having flat bottom with 250 nodes	68
Fig. 4.7 Flowchart for the grid generation	69-70
Fig. 5.1 Element types in one dimension	71
Fig. 5.2 Element types in two dimension with ϕ specified at the nodes	71
Fig. 5.3 Linear interpolation functions for two node line elements	72
Fig. 5.4 Quadratic interpolation functions for three node line elements	72

LIST OF FIGURES(continued)

Figure	Page
Fig. 5.5 Linear interpolation functions for a three node triangular element $\phi = N_1\phi_1 + N_2\phi_2 + N_3\phi_3$	73
Fig. 5.6 Assembled contributions $K\phi = Q$ (a) System matrices (b)Contributing elements	74
Fig. 6.1 (a) The conventional Hall plates (b) Partly deformed Hall plates (c) Semicircular Hall plates	75
Fig. 6.2 Device geometry for the simulation of the VHD	76
Fig. 6.3 Grid for the simulation of the VHD	76
Fig. 6.4 The Hall voltage vs. sensor probe positions (a) $\mu^*B = 0.01$ (b) $\mu^*B = 0.1$ (c) $\mu^*B = 0.2$	77
Fig. 6.5 Potential distribution at optimum position ($\mu^*B = 0.2$)	78
Fig. 6.6 Equipotential lines at optimum position ($\mu^*B = 0.2$)	79
Fig. 6.7 Equipotential and current lines from Huiser and Baltes	79
Fig. 6.8 The Hall voltage vs. magnetic induction	80
Fig. 6.9 Flowchart for the simulation of the VHD	81-83
Fig. 7.1 Schematic of an n-channel MOSFET	84
Fig. 7.2 Grid for the simulation of the MAGFET ($W = L = 100\ \mu m$)	84
Fig. 7.3 Top view of the double drain MAGFET	85
Fig. 7.4 Top view of the triple drain MAGFET	86

LIST OF FIGURES(continued)

Figure	Page
Fig. 7.5 Equipotential lines of the double drain MAGFET (μ*B = 0.1, W = L = 100 μm)	87
Fig. 7.6 Equipotential lines of the double drain MAGFET (μ*B = 0.2, W = L = 100 μm)	88
Fig. 7.7 Potential distribution of the double drain MAGFET (μ*B = 0.2, W = L = 100 μm)	89
Fig. 7.8 Equipotential lines of the triple drain MAGFET (μ*B = 0.2, W = L = 100 μm)	90
Fig. 7.9 Potential distributions of the triple drain MAGFET (μ*B = 0.2, W = L = 100 μm)	91
Fig. 7.10 Drain current imbalance vs. magnetic induction	92
Fig. 7.11 Flowchart for the simulation of the MAGFET	93-94

LIST OF SYMBOLS

Symbol	Name	Unit
$\vec{B}$	Magnetic induction vector	tesla
C_{ox}	Gate oxide capacitance per unit area	farad/cm^2
D_n	Electron diffusion constant	cm^2/sec
D_p	Hole diffusion constant	cm^2/sec
e	charge of a carrier ($\pm 1.6 \cdot 10^{-19}$)	coulomb
$\vec{E}$	Electric field vector	volt/cm
E_i	Intrinsic Fermi level	eV
E_{FP}	Bulk Fermi level	eV
$\vec{F}$	Lorentz force vector	newton
G_n	Electron generation rate	cm^{-3} sec^{-1}
G_p	Hole generation rate	cm^{-3} sec^{-1}
$\vec{H}$	Magnetic field strength vector	(volt-sec/cm^2)^2
I_D	Drain current	ampere
$\vec{J}_n$	Electron current density vector	amp/cm^2
$\vec{J}_p$	Hole current density vector	amp/cm^2
k	Boltzman's constant	Joule/^{o}K
L	Channel length	μm
n	Electron concentration	cm^{-3}
N_A^-	Ionized acceptor concentration	cm^{-3}
N_D^+	Ionized donor concentration	cm^{-3}

LIST OF SYMBOLS(continued)

Symbol	Name	Unit
p	Hole concentration	cm^{-3}
q	Magnitude of the electronic charge	coulomb
Q_n'	Surface charge density	$coulomb/cm^2$
R_j	Average resistance between adjacent point i and j	ohm
r_n	Hall factor for electron	
r_p	Hall factor for hole	
R_H	Hall coefficient	
R_n	Electron recombination rate	$cm^{-3}\ sec^{-1}$
R_p	Hole recombination rate	$cm^{-3}\ sec^{-1}$
T	Temperature	Kelvin
$\vec{v}$	Drift velocity of moving charge carrier	cm/sec
t_{ox}	Oxide thickness	nm
V_B	Substrate voltage	volt
V_D	Drain voltage	volt
V_{FB}	Flatband voltage	volt
V_G	Gate voltage	volt
V_S	Source voltage	volt

LIST OF SYMBOLS(continued)

Symbol	Name	Unit
V_T	Threshold voltage	volt
W	Width of channel	μm
μ_o	Permeability of vacuum	cm^2/volt-sec
μ_n	Electron drift mobility	cm^2/volt-sec
μ_p	Hole drift mobility	cm^2/volt-sec
μ_n^*	Electron Hall mobility	cm^2/volt-sec
μ_p^*	Hole Hall mobility	cm^2/volt-sec
ϵ	Permittivity	farad/cm
θ_{Hn}	Hall angle	degree
τ_n	Mean free time between carrier collision	sec
ψ	Electrostatic potential	volt
δ	Relative current imbalance between two outer drains	
Δ	Area of a triangle	μm^3
∇	Nabla operator	

1. INTRODUCTION

Since the bipolar transistor was announced in 1947, the semiconductor device field has grown rapidly. Coincident with the growth in semiconductor device development, the literature concerning semiconductor device and technology issues has literally exploded. The advent of Very Large Scale Integration (VLSI) has certainly revealed the need for a better understanding of basic device behavior. One of the characteristic features of classical (analytical) device modeling is the separation of the interior of the device into different regions, treated independently, with restrictive and sometimes drastic assumptions. The major breakthrough for VLSI was the miniaturization of the single transistor. The carrier transport effects (e.g. short channel effects) in such miniaturized transistor can not be explained by the simple classical models since severe or wrong assumptions lead to a conclusion far from the actual behavior of the device. As a consequence numerical simulation based on comparatively fundamental differential equations has become necessary and popular. This trend has been supported considerably by the enormous progress in the hardware and software of computers [1,2].

At the outset it seems necessary to clarify the frequently used terms such as *modeling*, *analysis* and *simulation*. According to Selberherr [2], *Modeling* is the description of something that can not be directly observed. For semiconductor devices, *modeling* is the mathematical description of the device, such as equations and boundary conditions, under the proper assumptions of the device under study. *Analysis* is the examination of something by separating the whole into its component parts. For semiconductor devices, *analysis* is an analytical examination by separating the device into several regions, where each region assumes a certain model. *Simulation* is the imitative representation of the functioning of one system or process by means of the functioning of another. For semiconductor devices, *simulation* is the numerical examination of the model by using computers.

The development of a new semiconductor device involves several design and fabrication cycles until a specified goal is reached. Device modeling can substantially reduce

the number of costly trial and error steps in this development. However, device modeling is not only driven by development cost reduction, but also by scientific curiosity. Device modeling provides a unique insight into the functioning of devices since it gives the distributions of various physical quantities in the interior of the device, a region which is not readily accessible for experimental measurements [3]. For example, device modeling seems the only way to explain the operating principles of the magnetotransistor, an area which has become recently the subject of disputes. Furthermore, computer resources are becoming cheaper as opposed to the rising costs of experimental investigations. Already many device modeling groups in the U.S.A. and Japan are using supercomputers to simulate complex models. Supporting software packages such as SPARSEPAK (for the calculation of pseudo inverse matrix), IMSL (mathematical program library) and DISSPLA (for graphic representation) are widely being used. Thus we expect that numerical modeling of devices will become more and more important in the near future. In the following paragraphs, application and variety of the magnetic-field sensitive devices which we are going to model, will be described.

In general, an electronic signal-processing system consists of three blocks [4] :

1. an input transducer.
2. a signal processor.
3. an output transducer such as a display, transmitter, actuator or memory.

In the input transducer, often called a sensor, a non-electrical measurand such as temperature, pressure, magnetic-field or chemical concentration is converted into an electrical signal. In the signal processor the electrical signal is in some way modified, i.e. amplified, filtered, or converted from an analogue to a digital signal. In the output transducer the electrical signal is converted into a signal which can be perceived by one of our senses (display devices). The signal can also be stored (memory devices), cause some action, e.g. closing a valve or printing a character (actuator) or be transmitted to another location (transmitters). For sensors it is sometimes necessary to use a third intermediary signal carrier in the

conversion. Transducers which make use of intermediary signal carriers are called 'tandem transducers'.

A Magnetic-Field Sensor (MFS) is an input transducer which is capable of converting $\vec{H}$ (magnetic field strength vector) into an electrical signal. If a nonmagnetic signal is measured by means of an intermediary conversion into $\vec{H}$, a magnetic-field sensor is also needed. Thus we can distinguish between direct and indirect MFS applications [4] :

1. In direct applications MFS is used as a magnetic-field measuring device (magnetometor).
   a. earth magnetic-field measurement.
   b. the reading of magnetic tapes or discs.
   c. magnetic apparatus control.
2. In indirect applications MFS is used as a tandem transducer.
   a. linear and angular displacement detection.
   b. potential-free detection.
   c. wattmeter.
   d. biomagnetometry.

There are several fundamental quantities in magnetic-field theory, such as $\vec{H}$ and $\vec{B}$ (magnetic flux density or induction vector). The relationship is $\vec{B} = \mu\mu_o\vec{H}$ with μ_o being the permeability of vacuum and μ being the relative permeability [3]. We can distinguish two classes of MFS. The first is the class of self-generating MFS involving high permeability materials ($\mu >> 1$). These transducers commonly employ one or more coils, but coils can not be implemented very easily in IC technology. Although there is a thin film (NiFe, Au) layer technology for self-generating MFS, the application of thin film layer technology for conductors is not favored by IC manufacturers [4]. The second is the class of non self-generating MFS involving low permeability material ($\mu \simeq 1$). The MFS in silicon makes use of the Lorentz force, which acts on moving charge carriers,

$$\vec{F} = e\vec{v} \times \vec{B} \ , \qquad (1.1)$$

where $\vec{v}$ is the drift velocity and e is the charge of the carrier.

Semiconductor materials such as Si, GaAs, InSb, or InAs can be used to make MFS. Since electron mobility is larger than hole mobility, the materials are commonly n-type. For a large Hall angle, which will be defined in Chapter 2, high electron mobility semiconductor materials such as InSb, InAs are favoured. However, these materials have several disadvantages. Since InSb and InAs have a small bandgap they have to be used at low temperatures to avoid any intrinsic behavior. Furthermore, there is no established batch fabrication in InSb or InAs. For GaAs it is questionable whether the performance/price ratio can match that of silicon because GaAs IC technology is still costly [3,4].

In this thesis the numerical simulation of semiconductor magnetic sensors such as the Vertical Hall effect Device (VHD) and the Magnetic-field sensitive Field Effect Transistor (MAGFET) is presented. All of the results are obtained using an AMDAHL 5860 running MTS. The programming language used is FORTRAN VS which is a dialect of ANSI FORTRAN 77 level language.

Chapter 2 explains the basic phenomena occuring in semiconductor devices in the presence of a magnetic field. Depending on the device geometry, doping and operating conditions the action of the Lorentz force on the moving charge carriers results in the Hall effect, carrier deflection and magnetoconcentration. Chapter 3 gives the basic equations governing semiconductor behavior and previous works on modeling. First it explains the equations when there is no magnetic field. If a magnetic field exists, then the equations become more complicated. Most of the modeling works have been done by the finite difference method. Modeling using the finite difference method and the lumped circuit cell method will be discussed.

Chapter 4 and Chapter 5 are devoted to the numerical and computational method. In Chapter 4, the adaptive grid generation method which is necessary for the efficient solution of the Partial Differential Equations (PDE) is described. Chapter 5 gives the mathematical procedures which discretize the Poisson equation when using the finite element method. This

equation governs the behavior of the devices to be modeled.

Chapter 6 and Chapter 7 are devoted to the actual simulation of specific devices such as the VHD and the MAFGET. Chapter 6 describes the geometrical concepts of the VHD and the numerical technique. Presented as well are the generated results and discussions. Chapter 7 describes the simulation of the MAGFET. The assumptions involved, boundary conditions and the numerical technique are described. Results and discussions are also included.

Finally Chapter 8 summarizes the above works and outlines the possible areas which require further investigations.

2. MAGNETIC FIELD EFFECTS IN SEMICONDUCTOR

The self-generating transduction, governed by Faraday's Law of induction, can be used with inductive components to detect temporal variations of the flux density. However, coils can not be implemented very easily in IC technology, and furthermore a serious drawback of this method is that static fields are not detected in a simple way. In contrast, the action of the Lorentz force on the moving charge carriers such as electrons and holes in nonmagnetic semiconductors may be used for the detection of static and dynamic flux densities. Several phenomena such as the Hall effect, carrier deflection, magnetoconcentration and magnetoresistance result as a consequence of the Lorentz force. These phenomena or a combination of them may be exploited in the design of MFS [4].

2.1 Hall Effect

If a magnetic field is applied perpendicularly to the direction of current density in a semiconductor slab, then the Lorentz force deflects the moving charge carriers in a direction perpendicular to both current density and magnetic field vectors, resulting in a transverse electric field across the slab (Fig.2.1).

If L>>W in the device shown in Fig. 2.2 [5], then the Lorentz force will be balanced by the force caused by the Hall electric field. As far as the net force on moving charge carriers is concerned there remains only the force due to the applied electric field. Thus current lines are parallel to the y-axis. However, the resulting electric field is composed of the Hall electric field and the applied electric field. The resulting electric field is not parallel to the y-axis but skewed by a Hall angle. Hence, the equipotential lines are also skewed by the same Hall angle with respect to the x-axis.

A device based on the Hall effect was first proposed by Gallagher and Corak [6]. This device was a p-channel MAGFET combined with Hall probes. The Hall probes were located on both sides of the channel.

Hirata and Suzuki [7] fabricated a similar MAGFET using NMOS technology. A sensitivity of 640 V/AT was obtained when the device was operated in depletion mode linear region. The MAGFET was used as the sensor part of a contactless keyboard.

2.2 Carrier Deflection

If W>>L in the device shown in Fig. 2.3 [5], then the Hall field makes little effect on the carriers in the bulk. In this case the current density vector is no longer parallel to the y-axis, but is rotated by the Hall angle. In fact this situation is complementary to the Hall effect, in which the current lines remain unaffected and the equipotential lines are rotated.

Usually double or triple drain MOS magnetic sensors depend on this effect. In these devices, the current imbalance between drains caused by carrier deflection is used as the measure of the magnetic field. Fry and Hoey [8] invented a p-channel MAGFET with split drains. Popović and Baltes [9] made a CMOS version of the MAGFET but with the drains cross coupled. Their device exhibited 10,000 V/AT, which was 10 times higher than that of the best known MAGFET and had offset-voltage control capability through gate contacts. The operating principles of the Magnetotransistors could also be explained in terms of carrier deflection in the base region [10,11].

2.3 Magnetoconcentration

If electrons and holes are simultaneously injected from opposite sides of an intrinsic semiconductor slab, then the Lorentz force deflects both types of carrier to the same surface. The Lorentz force is compensated by a carrier concentration gradient perpendicular to the magnetic and electric field vectors (Fig. 2.4).

The Magnetodiode is based on this effect. Pfleiderer [12] presented theoretical and experimental analysis of the magnetodiode. Kamarinos *et al.* [13] proposed a silicon on sapphire(SOS) magnetodiode which employs a Si-SiO_2 interface on top for obtaining a low recombination velocity and a Si-Al_2O_3 interface at the bottom for a high recombination

velocity.

2.4 Hall Angle

The Hall angle is defined as the angle between two electric fields [14]

$$\tan \Theta_{Hn} = \frac{E_y}{E_x} \quad . \tag{2.1}$$

However it can be defined as the angle between the current density and the resulting electric field.

$$\tan \Theta_{Hn} = \frac{|\vec{J} \times \vec{E}|}{|\vec{J} \cdot \vec{E}|} \quad . \tag{2.2}$$

2.5 Ideal Hall Voltage

The Hall voltage is given by [15]

$$V_y = \frac{E_y}{W} = \frac{R_H I_X B_z}{d} \quad , \tag{2.3}$$

where W is the width of the semiconductor slab, d is the sample thickness in the direction of the magnetic field (see Fig. 2.1) and R_H is the Hall coefficient.

In the weak field approximation ($(\mu^* B_z)^2 \leq 0.1$) R_H for the n-type semiconductor reads [15]

$$R_H = \frac{r_n}{ne} \quad , \tag{2.4}$$

where r_n is called the Hall factor for the electron and stands for

$$r_n = \frac{\langle \tau_n^2 \rangle}{\langle \tau_n \rangle^2} \quad , \tag{2.5}$$

$\langle\ \rangle$ denotes the expectation value. The parameter τ_n is the mean free time between carrier collisions. r_n is about 1.15 at room temperature for low donor concentration [4,16].

V_y for the n-type semiconductor now reads

$$V_y = \frac{r_n I_x B_z}{ned} \tag{2.6}$$

$$= \frac{\mu_n^*}{\mu_n} \cdot \frac{I_x B_z}{ned} = \frac{1}{ne} \cdot \frac{\mu_n^* B_z}{\mu_n} \cdot \frac{I_x}{d} .$$

If the magnetic induction B_z is not weak, then r_n and n are functions of B_z and so is R_H, hence V_y no longer has a linear relationship with B_z. Hall coefficient for the p-type semiconductor is r_p/pe and r_p $(=\mu_p^*/\mu_p)$ is about 0.7 at room temperature for low doping [4,16].

3. BASIC SEMICONDUCTOR EQUATIONS AND PREVIOUS MODELING WORKS

To simulate an arbitrary semiconductor structure for various operating conditions and physical parameters, it is necessary to construct a mathematical model. The equations which form this mathematical model are commonly called the basic semiconductor equations. The basic semiconductor equations consist of Poisson's equation, the continuity equations for electrons and holes and the current relations for electrons and holes. These equations are valid for the majority of engineering applications, particulary for Si devices [2]. Additional equations are needed to describe complicated effects such as thermoelectric or thermomagnetic effects in the semiconductor devices. However, these additional effects demand large computer resources and can pose a great deal of programming difficulties. Thus the equations to be used in simulation should be determined on the consideration of the performance/efficiency ratio.

3.1 Current Density Equations

3.1.1 Current density equations without magnetic field

The equations (3.1) - (3.6) are obtained under several assumptions such as zero magnetic induction, no time and spatial variations of carrier temperature, slowly varying impurity concentration, elastic scattering mechanism, infinitely large semiconductor etc. These assumptions are summarized on pp. 17 in Ref. [2].

The diffusion currents are proportional to the gradient of the carrier concentration, viz.,

$$\vec{J}_{n,diff} = qD_n \vec{\nabla} n \quad , \tag{3.1}$$

$$\vec{J}_{p,diff} = -qD_p \vec{\nabla} p \quad , \tag{3.2}$$

where D_n and D_p are diffusion constants for electrons and holes respectively and n and p are the electron and hole concentrations respectively and $q = 1.6 \cdot 10^{-19}$ as the magnitude of the electron charge. The drift currents are propotional to the electric field, viz.,

$$\vec{J}_{n,drift} = q\mu_n n\vec{E} \quad , \tag{3.3}$$

$$\vec{J}_{p,drift} = q\mu_p p\vec{E} \quad . \tag{3.4}$$

The total conduction currents are the sum of the diffusion and drift current components,

$$\vec{J}_n = q\mu_n n\vec{E} + qD_n\vec{\nabla}n \quad , \tag{3.5}$$

$$\vec{J}_p = q\mu_p p\vec{E} - qD_p\vec{\nabla}p \quad . \tag{3.6}$$

3.1.2 Current density equations with magnetic field

Under the influence of the magnetic field current density equations should be modified as follows [17] ;

$$\vec{J}_n = (q\mu_n n\vec{E} + qD_n\vec{\nabla}n) - \mu_n^* (\vec{J}_n \times \vec{B}) \quad , \tag{3.7}$$

$$\vec{J}_p = (q\mu_p p\vec{E} - qD_p\vec{\nabla}p) + \mu_p^* (\vec{J}_p \times \vec{B}) \quad , \tag{3.8}$$

where the second term of the RHS comes from Lorentz force. These equations correspond to the leading terms of a weak field expansion and are valid when $(\mu^* B)^2 \leq 0.1$ [3,17]. Let us consider just the electron current density equation (3.7) and to simplify the notation [5], let

$$\vec{A}_n = q\mu_n n\vec{E} + qD_n\vec{\nabla}n \quad . \tag{3.9}$$

If we take the vector product of $\vec{B}$ with (3.7), then

$$\vec{B} \times \vec{J}_n = \vec{B} \times \vec{A}_n - \mu_n^* [\, |\vec{B}|^2 \vec{J}_n - (\vec{B} \cdot \vec{J}_n)\vec{B} \,] \quad . \tag{3.10}$$

If we take the dot product of $\vec{B}$ with (3.7), then

$$\vec{B} \cdot \vec{J}_n = \vec{B} \cdot \vec{A}_n \quad . \tag{3.11}$$

By applying (3.10) and (3.11) to (3.7) the electron current density equation now becomes,

$$\vec{J}_n = \vec{A}_n + \mu_n^* \{ \vec{B} \times \vec{A}_n - \mu_n^* [\, |\vec{B}|^2 \vec{J}_n - (\vec{B} \cdot \vec{A}_n)\vec{B} \,] \} \quad . \tag{3.12}$$

Making an explicit form of $\vec{J}_n$ results in

$$\vec{J}_n = \frac{1}{1 + (\mu_n^* |\vec{B}|)^2} [\, \vec{A}_n + \mu_n^* (\vec{B} \times \vec{A}_n) - (\mu_n^*)^2 (\vec{B} \cdot \vec{A}_n)\vec{B} \,] \quad . \tag{3.13}$$

Equation (3.13) is the general expression relating the current density $\vec{J}_n$ and the magnetic induction $\vec{B}$ [5]. This equation comprises the isothermal galvanomagnetic effects for electrons. It accounts for the direct effects of temperature on carrier concentration, diffusion, and mobility, but does not include thermomagnetic or thermoelectric effects [3].

We can divide (3.13) into two cases depending on the relative direction of the $\vec{J}_n$ and $\vec{B}$ [5]. When $\vec{J}_n$ is perpendicular to $\vec{B}$, (3.11) becomes

$$\vec{B} \cdot \vec{J}_n = \vec{B} \cdot \vec{A}_n = 0 \tag{3.14}$$

and substituting (3.14) into (3.13) yields,

$$\vec{J}_n^{\perp} = \frac{1}{1 + (\mu_n^* |\vec{B}|)^2} [\, \vec{A}_n + \mu_n^* (\vec{B} \times \vec{A}_n) \,] \tag{3.15}$$

When $\vec{J}_n$ is parallel to $\vec{B}$, from (3.7) we obtain

$$\vec{J}_n^{\,II} = \vec{A}_n \ . \qquad (3.16)$$

3.1.3 Current density components with magnetic field

If $\vec{B}$ is chosen perpendicular to the device plane and parallel to the z-axis, i.e. $\vec{B} = (0,0,B_z)$, from (3.15) the electron current density components in the device plane can be expressed as follows [5] ;

$$J_{nx} = \frac{1}{1 + (\mu_n^* B_z)^2} \left[\left(q\mu_n n E_x + qD_n \frac{\partial n}{\partial x}\right) - \mu_n^* B_z \left(q\mu_n n E_y + qD_n \frac{\partial n}{\partial y}\right) \right] , \qquad (3.17)$$

$$J_{ny} = \frac{1}{1 + (\mu_n^* B_z)^2} \left[\left(q\mu_n n E_y + qD_n \frac{\partial n}{\partial y}\right) + \mu_n^* B_z \left(q\mu_n n E_x + qD_n \frac{\partial n}{\partial x}\right) \right] , \qquad (3.18)$$

$$J_{nz} = \left(q\mu_n n E_z + qD_n \frac{\partial n}{\partial z}\right) . \qquad (3.19)$$

The hole current density components can also be expressed in the same way as follows ;

$$J_{px} = \frac{1}{1 + (\mu_p^* B_z)^2} \left[\left(q\mu_p p E_x - qD_p \frac{\partial p}{\partial x}\right) + \mu_p^* B_z \left(q\mu_p p E_y - qD_p \frac{\partial p}{\partial y}\right) \right] , \qquad (3.20)$$

$$J_{py} = \frac{1}{1 + (\mu_p^* B_z)^2} \left[\left(q\mu_p p E_y - qD_p \frac{\partial p}{\partial y}\right) - \mu_p^* B_z \left(q\mu_p p E_x - qD_p \frac{\partial p}{\partial x}\right) \right] , \qquad (3.21)$$

$$J_{pz} = (q\mu_p pE_z - qD_p \frac{\partial p}{\partial z}) \quad . \tag{3.22}$$

3.2 Current Continuity Equations

The current continuity equations are a statement of the conservation of the charge within a semiconductor [18]. It means that sources and sinks of the total conduction current are fully compensated by the time variation of the mobile charge.

$$\frac{\partial n}{\partial t} = \frac{1}{q} \vec{\nabla} \cdot \vec{J}_n + G_n - R_n \quad , \tag{3.23}$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q} \vec{\nabla} \cdot \vec{J}_p + G_p - R_p \quad . \tag{3.24}$$

In the derivation we assume all charges in the semiconductor, except the mobile carriers such as electrons and holes, are time invariant. Thus, we neglect the influence of charged defects, e.g. vacancies, dislocations, deep recombination traps, which may change their charge state in time [2]. Under the steady state condition (3.23) and (3.24) become

$$\vec{\nabla} \cdot \vec{J}_n = -q (G_n - R_n) \quad , \tag{3.25}$$

$$\vec{\nabla} \cdot \vec{J}_p = q (G_p - R_p) \quad . \tag{3.26}$$

3.3 Poisson's Equation

The relationship between the space charge density and the electrostatic potential is described by Poisson's equation [18].

$$\vec{\nabla} \cdot (\varepsilon \vec{\nabla} \psi) = -q \, (N_D^{+} - N_A^{-} + p - n) \quad , \tag{3.27}$$

where ψ is the electrostatic potential and N_D^{+}, N_A^{-} are the ionized donor and acceptor concentrations respectively. The permittivity ϵ may be treated as a scalar quantity.

3.4 Modeling by Lumped Circuit Cell Method

Arnold [19] simulated the Hall effect in an inhomogeneous inversion layer by modeling the inversion layer as a two dimensional resistor network.

The condition of current continuity in the inhomogeneous conductor where the conductivity varies with position may be written as a finite difference equation on a two dimensional mesh of nodes.

$$\sum_j (V_i - V_j) / R_j = 0 \quad , \tag{3.28}$$

where V_i is the potential at the point i and $1/R_j$ is the average conductance between adjacent points i and j. The above equation is equivalent to the Kirchhoff current law for a regular resistor network.

From the definition of the Hall angle,

$$\tan \theta_H = \frac{E_y}{E_x} = \frac{V_y}{V_x} \cdot \frac{d_x}{d_y} \quad . \tag{3.29}$$

Thus,

$$V_y = (\tan \theta_H)\, V_x \frac{d_y}{d_x} = \mu^* B\, V_x \frac{d_y}{d_x} \quad . \tag{3.30}$$

In a square cell, $d_y = d_x$ and the above equation now reads

$$V_y = \mu^* B\, V_x \quad , \tag{3.31}$$

where V_x is the voltage drop across the resistor in the x direction and V_y is the local Hall voltage, that is the voltage drop across resistor in y direction. Using the initial values of voltage, we solve (3.28) to get the new V_i and solve (3.31) to get the new V_y. In this way we solve (3.28) and (3.31) until convergence of the local Hall voltage V_y is reached.

3.5 Modeling by the Finite Difference Method

Numerical device modeling using finite difference methods is highly developed for zero magnetic field [20-24]. Finite difference methods are well established in terms of discretizing the equations and error analysis. It is relatively easy to incorporate boundary conditions in finite difference methods. The modeling of a magnetic field sensitive semiconductor device is made particularly difficult by the fact that the magnetic field introduces some asymmetry into the current equations and boundary conditions [25].

Zieren [4] simulated the Vertical Magnetotransistor (VMT) using a finite difference method by neglecting space charge and magnetoconcentration effects. He solved Laplace's equation using a five node discretization scheme and obtained potential distributions in the epitaxial region of the device. The voltage V_0 on node 0 may be approximated by

$$V_0 = (V_1 + V_2 + V_3 + V_4) \;/\; 4 \quad , \tag{3.32}$$

where 1, 2, 3, 4 are adjacent nodes in the rectangular grid.

The University of Alberta sensor modeling group (ALBERTINA) first presented numerical solutions of the coupled nonlinear elliptic Partial Differential Equations (PDE) used to model the carrier transport in the magnetic sensitive semiconductor devices [25-32]. They discretized the continuity equations using a finite box type procedure, taking into account the Scharfetter-Gummel approximation [33]. After discretization, three equations were linearized by a Newton-iteration scheme. However, the linear system obtained was, in many cases of practical interest, unsuited for an iterative solution because of the numerical values the out-of-diagonal elements took on. This problem was solved by generalizing the well-known Scharfetter-Gummel scheme for the case of two-dimensions and nonzero magnetic field. The solutions were obtained using a nonuniform rectangular grid and the Successive Line OverRelaxation (SLOR) iteration scheme with a relaxation factor between 1.0 and 1.9 depending on the device parameters and the initial guess values [27].

In the doped materials, depending on the device geometry either a Hall effect or carrier deflection dominated. In the case of doped materials, the results from the ALBERTINA modeling program agrees with prior results obtained by simple analytical models or conformal mapping techniques [27]. However, if the doping level is of the same order as the intrinsic concentration and if significant electric and magnetic fields are present, then the space charge can not be neglected. This is clearly demonstrated for intrinsic or nearly intrinsic Hall plates and PIN-diodes where magnetoconcentration effects prevail [27,30].

Split-drain MAGFET devices operating in the linear region were modeled using the same numerical procedures. Current density and carrier concentration were integrated over the channel depth to obtain a two-dimensional problem. The current imbalance between the two outer drains produced by the (perpendicular) magnetic induction was demonstrated. The modeling results allow prediction of the sensitivity of split-drain MAGFET devices and optimization of device geometry. Combining the sensitivity results with current MOSFET noise models, the minimum detectable induction can be predicted as well [29,31,32].

3.6 Analytical Modeling

Semiconductor devices have been modeled analytically in many ways. One of the most popular techniques is the conformal mapping technique. Sugiyama [34] solved the two-dimensional Laplace's equation on a rectangular three-terminal Hall element by means of a Schwarz-Christoffel's conformal mapping. Equipotential lines and electric force lines were obtained as complex conjugates of each other.

Huiser, Baltes and Nathan [32,35,36] calculated the potential and current distribution in the VHD by means of Green's function. Again they used the conformal mapping technique to obtain a geometry with a better resemblance to Popović's configuration.

4. GRID GENERATION

4.1 Basic Concept

4.1.1 Comparison of the finite element method and the finite difference method

Several numerical methods have been developed over the years to solve partial differential equations. Among them the most commonly used are the finite difference method and the finite element method. In order to obtain a finite difference equations for the given PDE, the region to be examined is covered by a rectilinear grid which is composed of vertical and horizontal lines. The grid for the finite difference method will reasonably cover the entire region, but the boundaries are usually approximated roughly [37]. The finite element method requires dividing the domain of the solution into a finite number of simple subdomains, the finite elements, to construct an approximation of the solution over the collection of finite elements. The grid for the finite element method is more conveniently suited to the boundary shape [38]. Moreover, a locally refined grid can be used for the finite element method. Local refinement makes device simulation very efficient because with a given number of nodes more accurate results can be obtained depending on the node distribution. Furthermore, it may determine whether the program converges or not. This point will be discussed in detail in the Section 4.1.2.

Thus we can conclude that the finite element method is better suited for problems having complex geometrical shapes and can make the program more efficient [37]. To exploit the advantage of the finite element method we have to be able to generate a grid having nonrectangular elements and make refinements where necessary.

4.1.2 The reason for grid refinement

In the finite element modeling of semiconductor devices grid refinement is important for several reasons.

First, the unknown functions often exhibit sharp variations over a small piece of the domain, while they are almost constant in the remaining zones [39]. To save computer time we need a dense grid where the function changes rapidly and a coarse grid where the function changes slowly [38].

Second, before imposing constant boundary conditions the discretized partial differential equations yield an almost singular matrix [40]. If the constant boundary regions are small, the matrix after imposing the boundary conditions still remains almost singular. In this case the inverse matrix does not exist and any kind of numerical technique that involves a pseudo-inverse matrix will not work. For example, the VHD can not be numerically simulated without grid refinement because this device in the absence of a bulk contact shows constant boundary values only over a small part of the boundary. In such a case, a direct method such as Gauss Elimination produces significant roundoff error and loses significant figures in the pivots. An iterative method like Successive OverRelaxation(SOR) converges very slowly or does not converge at all. Moreover, a singularity of the matrix can occur when some of the rows are almost linearly dependent [41]. This kind of ill-conditioned matrix is impossible to detect in the practical sense before entering the solution procedure because the test of the condition would need the inverse, eigenvalue or determinant of the matrix, and usually these are just the quantities we wish to determine [41]. By locating a large number of nodes on the constant boundary region we can improve the condition of the matrix. In other words, more components in the matrix are explicitly determined by the constant boundary conditions and the matrix becomes less singular. From the author's experience, an average of six to seven times more nodes can be put by the grid refinement technique than the rectangular grid generation technique. Since the approximation error is bound by the powers of element size [40,42], the error in the dense grid region which is composed of small size elements is significantly reduced.

Finally, the very boundary conditions can impose interface or boundary singularities such as discontinuity of a prescribed value which give significant errors in the solution. These

errors can be reduced by a proper grid refinement technique which generates a denser grid near the singular points [40].

4.2 Grid Generation Procedure

4.2.1 Initial grid generation

We distinguish three types of polygon shapes (they have to be always within the first quadrant).

1. Rectangular shape : Sides of the rectangle are on x-y axes. Embedded domains are not allowed.
2. Flat bottom shape : Bottom side is on x-axis, and clockwise successive boundary points, starting from bottom left, never have decreasing x coordinate (i.e. no horizontal indents). Embedded domains are not allowed.
3. Any arbitrary shape is approximated by polygons.

Depending on the shape of the device, we use different ways to generate the initial grid.

1. Rectangular shape : The user has to specify the number of horizontal lines (h_1+1) and vertical lines (k_1+1). The nodes are generated by the intersection of horizontal lines and vertical lines. Then the triangular elements are generated.
2. Flat bottom shape : The user has to provide the coordinates of boundary points (clockwise from bottom left) and the number of horizontal lines (including the bottom line and the farthest point(s)). The scale of coordinates is one micron (1 μm). The user then is prompted for the specification of additional horizontal lines. The intersections of the horizontal lines with the boundary lines define the vertical lines. The program adds vertical and horizontal lines on boundary points not included in the generation scheme so far. The user then is prompted for the specification of additional vertical lines intersecting with horizontal boundary lines. Based on information about boundary, horizontal and vertical lines, the nodes and elements are generated.

3. Arbitrary shape : The user has to generate an initial grid by specifying the coordinates and node number of each node which belongs to a triangular element. In this case there is no restriction on the geometrical shape of the device because curved lines are approximated by polygon lines. However, there are some limitations. The user draws imaginary vertical and horizontal lines over the device domain. At least each boundary point has to be crossed by a vertical and a horizontal line. Then the user decides on the elements. The elements should be either right angled triangles or isosceles triangles. In other words, the element shapes should fit into the configuration class of Fig. 4.2 (a), (b), (c) and (d) including all their mirrored representations (along both axes).

4.2.2 Numbering scheme for an initial grid

For the initial grid, a systematic way of numbering is required. We need two kinds of numbering, i.e., the numbering of the nodes and of the elements. We also need to know which nodes belong to an element. This information is required for later use in the refinement step. In our program we assume that node numbers start from 0 and element numbers start from 1 (Fig. 4.1).

For a rectangular shape, The relationship between node numbering and element numbering is described by the algorithm that follows ;

```
for IY ← 1 to h1
    begin
        for IX ← 1 to k1
            begin
                J ← (IY-1)•k1 + IX ;
                I ← (IY-1)•(k1+1) + IX ;
                E(2J-1,1) ← I-1 ;
                E(2J-1,2) ← I ;
```

```
            E(2J-1,3) ← I + k1 + 1 ;
            E(2J,1) ← I + k1 + 1 ;
            E(2J,2) ← I + k1 ;
            E(2J,3) ← I-1 ;
        end ;
    end ;
```

Here, IY and IX are index variables for looping, J and I are index variables for array E which gives information regarding node numbering and element numbering. Other kinds of numbering schemes are also possible. During the refinement procedure any systematic way of numbering is destroyed; and since an initial grid contains a relatively small number of nodes compared to the final grid, any particular choice of numbering scheme (for an initial grid) does not have a significant effect on the efficiency of the program. For a nonrectangular shape, the elements which are located at the outside of the device region will be excluded after using above algorithm.

4.2.3 Grid refinement

After generating an initial grid, the program control goes to the grid refinement procedure. Our procedure involves two different ways of grid refinement. First, to refine the initial grid, the user has to provide the refinement criteria in terms of the gradient of some weighting function. If the gradient of the weighting function between any two nodes is greater than the value presented by refinement criteria, the refinement is performed. In semiconductor device modeling, the functional behavior of any important variable such as a doping profile can be a weighting function. There are two passes, viz., such as vertical direction refinement pass and horizontal direction refinement pass. After the refinement step between two nodes, the gradient between the next two nodes, either lying horizontally or vertically depending on the pass, is checked.

Second, the user can specify rectangular sub-regions to be refined and the scale of refinement (e.g. 5 times for finer grid) in these sub-regions. If the scale of refinement is greater than one, then only these sub-regions are refined regardless of the gradient of the weighting function and any other region will not be refined. The scale of refinement is reduced by 1 after one vertical and horizontal refinement pass. If the number of refinement is less than 1 (i.e. zero), then sub-regions are refined depending on the gradient of the weighting function, but any other region will not be refined.

During the refinement, the systematic numbering scheme generated in the initial grid generation step will be not be conserved. However, the correctly matched relationship between new node numbers and element numbers is updated and maintained. This is essential because the finite element procedure requires information about the nodes that belong to an element as well as the coordinates of the nodes.

4.2.3.1 Horizontal refinement pass

The horizontal refinement pass starts with the first element given by the numbering scheme. The program first checks to see if the element has a refinable shape. The refinable element shapes for horizontal refinement pass and the corresponding refined shapes are shown in Fig. 4.2 (a), (b), (c) and (d). There are other possible element shapes than those shown. However, we choose the shapes in Fig. 4.2 (a), (b), (c) and (d) as the refinable element shapes because we want to maintain regular structure for successive refinement in a simple way. If the element has a refinable shape and the gradient of the weighting function between two nodes lying in the horizontal direction is greater than the refinement criteria, then a new node is generated at the middle of the line connecting these two nodes, otherwise the element is skipped. In the cases shown in Fig. 4.2 (a) and the first case in Fig. 4.2 (d), only one new node is generated at the middle of the horizontal line. It should be noted that the first four cases in Fig. 4.2 (a) result in a same refined shape.

Slanted boundaries may occur when the device geometry is not rectangular. The cases in Fig. 4.2 (b), (c) and last two cases in Fig. 4.2 (d) show the refinement for the elements having slanted boundaries. In these cases the number of newly generated nodes is more than one. The dotted lines for the last two cases indicate that the shapes for these parts are not fixed.

The second element is then checked and possibly refined. All remaining elements are successively processed in this manner. Finally the control returns to the vertical refinement pass.

4.2.3.2 Vertical refinement pass

The same checking procedure is used as in the case of horizontal refinement, the difference being, the elements are now analyzed in the vertical direction.

4.2.4 Criteria to exit from grid generation

There are essentially two types of termination criteria. Both require change of the program. The grid refinement is terminated when the maximum number of nodes or the maximum number of elements is reached. In this way, one controls precisely the maximum number of nodes or the maximum number of elements. The maximum number of nodes and elements are interactively specified by the user. In another case the program stops only after a complete horizontal or vertical pass. In that case the maximum number of nodes and elements can not be precisely controlled by the user. From the author's experience, the latter case is usually desirable because the precise control of the maximum number of nodes or elements is not necessary in most cases of device modeling.

If the user provides values for refinement criteria which are higher than the gradient of the weighting function between any two nodes, it is possible that the refinement does not occur. To avoid such situations, refinement criteria are adaptively modified until refinement can proceed. After the horizontal or vertical refinement pass, the criteria given by the user is modified by a proper factor until the refinement takes place.

After the refinement step the program generates a picture of the final grid and asks the user if the generated grid is a good one. If the user answers 'NO' the grid generation will be repeated from the very beginning otherwise the control goes to the finite element procedure. A plot of the grid can be generated at each stage of the refinement. This is useful in the development period of finding a proper grid.

4.3 Results and discussions

All of the results shown in this chapter can be obtained in less than 20 CPU seconds. Figure 4.3 (a) shows an initial grid for the device with size 150 μm by 150 μm. All of the vertical lines and horizontal lines are equally spaced. For example, consider the weighting function

$$W(x,y) = |10^4 - x^2 - y^2| + 10^4 - x^2 - y^2 . \qquad (4,1)$$

The program then enters the refinement pass. The final grid, which has 400 nodes, is shown in Fig. 4.3 (b). The value of function W is zero if $x^2 + y^2 \geq 10^4$. Thus the grid is dense in the circle with center at the origin and radius 10^2. At the outside of this circle, the grid is very coarse.

Figure 4.4 (a) shows an initial grid for the device shape having a flat bottom and slanted boundaries. The maximum size in x direction is 100 μm and the maximum size in y direction is 60 μm. The rectangular sub-region, in which x is from 0 μm to 100 μm and y is from 50 μm to 60 μm, is chosen for refinement. The final grid having 200 nodes is shown in Fig. 4.4 (b). It is easily noticed that the grid inside the rectangular sbu-region is very dense. Figure 4.5 (a) is the second example of an initial grid for the device shape having a flat bottom and slanted boundaries. The maximum size in x direction is 100 μm and the maximum size in y direction is 60 μm. The rectangular sub-region, in which x is from 0 μm to 100 μm and y is from 50 μm to 60 μm is also introduced for refinement. The final grid, which has 230 nodes is shown in Fig. 4.5 (b). It is also easily noticed that the grid inside the rectangular

sub-region is very dense. Figure 4.6 provides another example of the grid refinement. The maximum size of the domain is 100 μm by 100 μm. The sub-region is from 60 μm to 100 μm in x direction and from 30 μm to 80 μm in y direction. The number of nodes is 250.

It should be noted that elements located near the sub-region are also refined. These elements act as the transition region between the sub-region and the rest of the device region.

There are several important considerations in grid generation.

1. The initial grid generation for the nonrectangular geometry : If the user specifies horizontal lines or vertical lines, then these lines will be reflected as vertical or horizontal lines respectively by the slanted boundaries. If all of the boundaries are slanted, then there is a possibility that the reflection process never ends. To reduce this kind of complexity we consider three different ways to generate an initial grid. The closer we get to arbitrary shape, the more information the user has to provide.
2. The refinement for the nonrectangular geometry : Curved lines may be approximated by polygon lines. Since we can refine elements having slanted boundaries, we can therefore refine any kind of device geometry. The idea chosen is based on consideration such as minizing the impact of refinement of an element to the other elements. The way chosen is shown in Fig. 4.2 (b), (c) and (d).
3. The obtuse angled elements : By careful selection of an initial grid, the number of obtuse angled elements can be maintained at few percents of the total number of elements. The obtuse angled element is important because this element can destroy the diagonal dominance property of the matrix resulted from the finite element procedure [43].

5. FINITE ELEMENT METHOD

5.1 Introduction

The finite element method has several ways of formulating finite element equations which describe the properties of the unknown variable ϕ over each element. One of the most popular approaches is the variational approach. For the variational approach the functional $I(\phi)$, whose minimum or maximum gives the solution we seek, has to be formulated. Thus we should try to find the nodal values of ϕ which make the functional $I(\phi)$ stationary. However, the problem is that the functional does not always exist and it is not easy to find this functional even if it does exist. Fortunately there is another approach which does not need a functional.

The weighted residuals approach is more a versatile approach than the variational counterpart. The weighted residuals approach begins with the governing equations of the problem and proceeds without relying on a functional. This approach is advantageous because it thereby becomes possible to extend the finite element method to problems where no functional is available. Applying the method of weighted residuals involves basically two steps. The first step is to assume the general functional behavior of the dependent variables in some way so as to approximately satisfy the given differential equation and boundary conditions. The second step is to solve the equations resulting from the first step and thereby specialize the general functional form to a particular function, which then becomes the approximate solution sought [37].

For example, we consider a typical problem. Suppose we want to find an approximate functional representation for a field variable ϕ governed by the differential equation [37]

$$L(\phi) - f = 0. \tag{5.1}$$

We approximate ϕ by u

$$\phi = u = \sum_{i=1}^{m} N_i C_i \quad , \tag{5.2}$$

where the N_i are the assumed functions, the C_i are either the unknown parameters or unknown functions of one of the independent variables and the upper limit m is the number of unknowns C_i.

When ϕ is substituted into (5.1) error is introduced,

$$L(u) - f = R \neq 0. \tag{5.3}$$

The method of weighted residuals seeks to determine the m unknowns, C_i, in such a way that the error R over the entire solution domain is small. This is accomplished by forming a weighted average of the error and specifying that the weighted average vanishes over the solution domain,

$$\int_D [\, L(u) - f \,] W_i \, dD = \int_D R W_i \, dD = 0 \tag{5.4}$$

$$(i = 1, 2, \cdots\cdots , m)$$

The most often used method to derive the finite element equations is known as Galerkin's method. According to the Bubnov-Galerkin method, the weighting functions are chosen to be the same as the approximating functions used to represent ϕ, that is $W_i = N_i$ [37]. Thus Galerkin's method requires that

$$\int_D [\, L(u) - f \,] N_i \, dD = 0 \tag{5.5}$$

$$(i = 1, 2, \cdots\cdots , m)$$

In the preceding discussion we assumed that we were dealing with the entire solution domain. However, because (5.1) holds for any point in the solution domain, it also holds for any collection of points defining an arbitrary subdomain or element of the whole domain. Thus we can formulate local approximation analogous to (5.2) for individual elements. Then

from Galerkin's method we can write the equations governing the behavior of an element as

$$\int_{D^{(e)}} [\, L(u^{(e)}) - f^{(e)} \,] \, N_i^{(e)} \, dD^{(e)} = 0 \qquad (5.6)$$

$$(i = 1, 2, \cdots\cdots, r)$$

where r is the number of the unknown parameters assigned to the element and the superscript (e) means that the variable is effective in a element.

However, we generally use ϕ instead of u in the finite element equations so (5.6) becomes ;

$$\int_{D^{(e)}} [\, L(\phi^{(e)}) - f^{(e)} \,] \, N_i^{(e)} \, dD^{(e)} = 0 \qquad (5.7)$$

$$(i = 1, 2, \cdots\cdots, r)$$

5.2 Finite Element Method for Poisson's Equation

Poisson's equation in two dimensional space is

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = f(\phi) \qquad (5.8)$$

with the boundary conditions such as

$$\phi = C \quad \text{on } \Sigma_1 \,, \qquad (5.9)$$

$$\left[\frac{\partial \phi}{\partial x} - a \frac{\partial \phi}{\partial y} \,,\, \frac{\partial \phi}{\partial y} + a \frac{\partial \phi}{\partial x}\right] \cdot \vec{n} = 0 \quad \text{on } \Sigma_2 \,, \qquad (5.10)$$

where a is a constant and $\vec{n}$ is the normal vector to the boundary Σ_2.

When we compare (5.1) and (5.8) we see that

$$L(\phi) = \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} \,. \qquad (5.11)$$

The finite element equations by Bubnov-Galerkin's method are (5.7). Superscript (e) can be deleted under the assumption that every equation is effective in the element if there is no specified domain. From now on every equation is valid in the element. We can modify finite element equations (5.7) to include natural boundary conditions;

$$\int_D [\, L(\phi) - \frac{\partial}{\partial x}\left(a \frac{\partial \phi}{\partial y}\right) + \frac{\partial}{\partial y}\left(a \frac{\partial \phi}{\partial x}\right) - f \,] N_i \, dD \tag{5.12}$$

$$= \int_D [\, \frac{\partial^2 \phi}{\partial x^2} - \frac{\partial}{\partial x}\left(a \frac{\partial \phi}{\partial y}\right) + \frac{\partial}{\partial y}\left(a \frac{\partial \phi}{\partial x}\right) + \frac{\partial^2 \phi}{\partial y^2} - f \,] N_i \, dx \, dy$$

$$= 0 \, .$$

To reduce the order of derivatives and introduce natural boundry conditions we use integration by parts [37],

$$\int_D N \, (\vec{\nabla} \cdot \vec{h}) \, dD = \int_\Sigma N \, (\vec{h} \cdot \vec{n}) \, d\Sigma - \int_D \vec{h} \cdot \vec{\nabla} N \, dD \, . \tag{5.13}$$

If we let,

$$\vec{h} = [\, \frac{\partial \phi}{\partial x} - a \frac{\partial \phi}{\partial y} \,] \vec{i} + [\, \frac{\partial \phi}{\partial y} + a \frac{\partial \phi}{\partial x} \,] \vec{j} \, , \tag{5.14}$$

$$N = N_i \, , \tag{5.15}$$

(5.12) now reads

$$\int_\Sigma N_i \, [\, (\frac{\partial \phi}{\partial x} - a \frac{\partial \phi}{\partial y}) \, n_x + (\frac{\partial \phi}{\partial y} + a \frac{\partial \phi}{\partial x}) \, n_y \,] \, d\Sigma \tag{5.16}$$

$$= \int_D [\, (\frac{\partial \phi}{\partial x} - a \frac{\partial \phi}{\partial y}) \frac{\partial N_i}{\partial x} + (\frac{\partial \phi}{\partial y} + a \frac{\partial \phi}{\partial x}) \frac{\partial N_i}{\partial y} \,] \, dx \, dy$$

$$+ \int_D f N_i \, dx \, dy \, .$$

The LHS is cancelled inside the global domain and is zero along the natural boundary due to the boundary condition. Thus (5.16) becomes ;

$$\int_D \left[\left(\frac{\partial \phi}{\partial x} - a \frac{\partial \phi}{\partial y} \right) \frac{\partial N_i}{\partial x} + \left(\frac{\partial \phi}{\partial y} + a \frac{\partial \phi}{\partial x} \right) \frac{\partial N_i}{\partial y} \right] dx \, dy \qquad (5.17)$$

$$= - \int_D f N_i \, dx \, dy \, .$$

5.2.1 Interpolation functions

To calculate (5.17) we have to specialize the form of N_i which is called the interpolation or base function. The requirement we place on the choice of the interpolation functions stem from the need to ensure that our approximate solution converges to the correct solution when we use an increasing number of smaller elements. To identify quantitively the admissible functions, we define the admissible class $C^m(D)$ [44].

Definition :

Suppose that D is a bounded region in R^3 the three dimensional space and that $f = f(x,y,z)$ is a given real-valued function of position in D. Then f is said to be of class C^m on D if f and all of its partial derivatives of order less than or equal to m are continuous at every point(x,y,z) in D, where m is a nonnegative integer.

For the second-order partial differential problems such as Poisson's equation, class C^0 interpolation functions are needed. For the fourth-order problems C^1 interpolation functions are needed [44].

Since our problem is to solve Poisson's equation we are interested in C^0 functions which are continuous throughout the global domain but their derivatives are not continuous. Figure 5.1 and Fig. 5.2 show the variety of element types [37]. For C^0 functions the unknowns, which are called *degrees of freedom*, assigned to an element are the values of ϕ on each node. For C^1 functions degrees of freedom may include the derivative of the nodal variable. Depending on the assumed functional behavior of ϕ, there are many choices for C^0 functions. In a one dimensional case if we think ϕ varies linearly we can use linear elements which have two nodes. With the values of ϕ on these two nodes we can interpolate ϕ uniquely over the element. Thus we can say that the choice of interpolation functions determine the element type or vise versa.

Figure 5.3 shows actual assemblage of the linear interpolation functions to represent the variable ϕ over one element in one dimension so that it gives [37]

$$\phi = \frac{x_2 - x}{x_2 - x_1}\phi_1 + \frac{x - x_1}{x_2 - x_1}\phi_2 \qquad (5.18)$$

$$= N_1(x)\phi_1 + N_2(x)\phi_2 \ .$$

The piecewise linear interpolation functions over the whole domain in one dimension can be as follows [45];

$$N_1(x) = \begin{cases} \dfrac{x_2 - x}{x_2 - x_1} & (x_1 \leq x \leq x_2) \\ 0 & (x_2 \leq x \leq x_n) \ , \end{cases} \qquad (5.19)$$

$$N_i(x) = \begin{cases} 0 & (x_1 \leq x \leq x_{i-1}) \\ \dfrac{x - x_{i-1}}{x_i - x_{i-1}} & (x_{i-1} \leq x \leq x_i) \\ \dfrac{x_{i+1} - x}{x_{i+1} - x_i} & (x_i \leq x \leq x_{i+1}) \\ 0 & (x_{i+1} \leq x \leq x_n) \ , \end{cases}$$

$$N_n(x) = \begin{cases} 0 & (x_0 \leq x \leq x_{n-1}) \\ \dfrac{x - x_{n-1}}{x_n - x_{n-1}} & (x_{n-1} \leq x \leq x_n) \end{cases},$$

where the coordinates of the nodes are denoted x_i and n is the number of nodes over the domain.

If we think the unknown variable varies quadratically, then we should use a three node element. Figure 5.4 illustrates the quadratic interpolation functions in one dimension [37].

Figure 5.2 shows the variety of element types in two dimensions and Fig. 5.5 gives the interpolation function of linear elements over a triangular element in two dimensions [46]. The order of the interpolation polynomial is determined by the number of degrees of freedom over the element. In other words, the number of coefficients in the polynomial should equal the number of nodal variables available to evaluate these coefficients. For the piecewise linear representation of a field variable over a triangle, we write the linear polynomial series as [37]

$$\phi = a_1 + a_2 x + a_3 y \quad . \tag{5.20}$$

The three nodal values of ϕ are then sufficient to find three coefficients. The coefficients a_i in a polynomial series represention of a field variable are called the *generalized coordinates* of the element. The usual procedure is to use the same number of generalized coordinates and degrees of freedom.

For integration in (5.17) we can use a numerical integration technique. However, if we use linear triangular elements, which allow ϕ to vary linearly over each element, we can evaluate the integration analytically. In the RHS of (5.17) if f is constant within an element, we have, from the integration formula of area coordinates (see Appendix),

$$\int_D f N_i \, dx \, dy = f \int_D N_i \, dx \, dy = f \frac{\Delta}{3} \quad , \tag{5.21}$$

where Δ is the area of the triangular element. If f is linear, we linearly interpolate f in terms of its nodal values (see Appendix),

$$f = f_1 N_1 + f_2 N_2 + f_3 N_3 \ , \tag{5.22}$$

$$Q_{1i} = \int_D f N_i \, dx \, dy = \int_D (f_1 N_1 N_i + f_2 N_2 N_i + f_3 N_3 N_i) \, dx \, dy \ ,$$

$$(i = 1, 2, 3)$$

$$[Q_1] = \frac{\Delta}{12} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} .$$

The LHS of (5.17) reads

$$\int_D \left[\left(\frac{\partial \phi}{\partial x}\right) \left(\frac{\partial N_i}{\partial x}\right) + a \left(\frac{\partial \phi}{\partial x}\right) \left(\frac{\partial N_i}{\partial y}\right) \right. \tag{5.23}$$

$$\left. - a \left(\frac{\partial \phi}{\partial y}\right) \left(\frac{\partial N_i}{\partial x}\right) + \left(\frac{\partial \phi}{\partial y}\right) \left(\frac{\partial N_i}{\partial y}\right) \right] dx \, dy \ ,$$

where

$$\phi = \sum_{i=1}^{3} N_i \phi_i \ ,$$

$$\frac{\partial \phi}{\partial x} = \begin{bmatrix} \frac{\partial N_1}{\partial x} & \frac{\partial N_2}{\partial x} & \frac{\partial N_3}{\partial x} \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} .$$

When i = 1, the above equation becomes

$$\begin{bmatrix} \int_D \frac{\partial N_1}{\partial x}\frac{\partial N_1}{\partial x}\,dx\,dy & \int_D \frac{\partial N_2}{\partial x}\frac{\partial N_1}{\partial x}\,dx\,dy & \int_D \frac{\partial N_3}{\partial x}\frac{\partial N_1}{\partial x}\,dx\,dy \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} \quad (5.24)$$

$$+ a \begin{bmatrix} \int_D \frac{\partial N_1}{\partial x}\frac{\partial N_1}{\partial y}\,dx\,dy & \int_D \frac{\partial N_2}{\partial x}\frac{\partial N_1}{\partial y}\,dx\,dy & \int_D \frac{\partial N_3}{\partial x}\frac{\partial N_1}{\partial y}\,dx\,dy \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix}$$

$$- a \begin{bmatrix} \int_D \frac{\partial N_1}{\partial y}\frac{\partial N_1}{\partial x}\,dx\,dy & \int_D \frac{\partial N_2}{\partial y}\frac{\partial N_1}{\partial x}\,dx\,dy & \int_D \frac{\partial N_3}{\partial y}\frac{\partial N_1}{\partial x}\,dx\,dy \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix}$$

$$+ \begin{bmatrix} \int_D \frac{\partial N_1}{\partial y}\frac{\partial N_1}{\partial y}\,dx\,dy & \int_D \frac{\partial N_2}{\partial y}\frac{\partial N_1}{\partial y}\,dx\,dy & \int_D \frac{\partial N_3}{\partial y}\frac{\partial N_1}{\partial y}\,dx\,dy \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} .$$

We can obtain other components when i = 2 and i = 3 in the same way.

From equ. (5.20) let us determine a_1, a_2 and a_3 by evaluating (5.20) at the nodes 1, 2 and 3 [37].

$$\phi_1 = a_1 + a_2 x_1 + a_3 y_1 , \quad (5.25)$$

$$\phi_2 = a_2 + a_2 x_2 + a_3 y_2 ,$$

$$\phi_3 = a_1 + a_2 x_3 + a_3 y_3 ,$$

and solving for a_1, a_2 and a_3 yields

$$a_1 = \frac{1}{2\Delta} [\phi_1(x_2 y_3 - y_2 x_3) + \phi_2(x_3 y_1 - x_1 y_3) + \phi_3(x_1 y_2 - x_2 y_1)] , \quad (5.26)$$

$$a_2 = \frac{1}{2\Delta} [\phi_1(y_2 - y_3) + \phi_2(y_3 - y_1) + \phi_3(y_1 - y_2)] ,$$

$$a_3 = \frac{1}{2\Delta} [\phi_1(x_3 - x_2) + \phi_2(x_1 - x_3) + \phi_3(x_2 - x_1)] ,$$

where Δ is the area of the triangle whose vertices are 1, 2 and 3.

Substituting (5.26) into (5.20) and rearranging terms give [37]

$$\begin{aligned} \phi &= \frac{1}{2\Delta} (\alpha_1 + \beta_1 x + \gamma_1 y)\phi_1 + \frac{1}{2\Delta} (\alpha_2 + \beta_2 x + \gamma_2 y)\phi_2 \\ &\quad + \frac{1}{2\Delta} (\alpha_3 + \beta_3 x + \gamma_3 y)\phi_3 \\ &= N_1\phi_1 + N_2\phi_2 + N_3\phi_3 \\ &= \sum_{i=1}^{3} N_i \phi_i , \end{aligned} \quad (5.27)$$

where

$$\alpha_1 = x_2 y_3 - x_3 y_2 , \quad (5.28)$$

$$\beta_1 = y_2 - y_3 ,$$

$$\gamma_1 = x_3 - x_2 .$$

The other coefficients are obtained through a cyclic permutation of the subcripts 1, 2 and 3.

N_i has the value 1 at the i'th node and the value 0 at the other two nodes.

By using (5.23) for the LHS of (5.17) and (5.22) for the RHS of (5.17), (5.17) becomes

$$\begin{bmatrix} (K_{11})_e & (K_{12})_e & (K_{13})_e \\ (K_{21})_e & (K_{22})_e & (K_{23})_e \\ (K_{31})_e & (K_{32})_e & (K_{33})_e \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} = - \begin{bmatrix} Q_{11} \\ Q_{12} \\ Q_{13} \end{bmatrix} . \tag{5.29}$$

Here,

$$(K_{11})_e = \int_D \frac{\partial N_1}{\partial x} \frac{\partial N_1}{\partial x} \, dx \, dy + a \int_D \frac{\partial N_1}{\partial x} \frac{\partial N_1}{\partial y} \, dx \, dy \tag{5.30}$$

$$- a \int_D \frac{\partial N_1}{\partial y} \frac{\partial N_1}{\partial x} \, dx \, dy + \int_D \frac{\partial N_1}{\partial y} \frac{\partial N_1}{\partial y} \, dx \, dy$$

$$= \frac{1}{4\Delta} [\beta_1^2 + \gamma_1^2] = \frac{1}{4\Delta} [(y_2 - y_3)^2 + (x_3 - x_2)^2] ,$$

$$(K_{12})_e = \int_D \frac{\partial N_2}{\partial x} \frac{\partial N_1}{\partial x} \, dx \, dy + a \int_D \frac{\partial N_2}{\partial x} \frac{\partial N_1}{\partial y} \, dx \, dy$$

$$- a \int_D \frac{\partial N_2}{\partial y} \frac{\partial N_1}{\partial x} \, dx \, dy + \int_D \frac{\partial N_2}{\partial y} \frac{\partial N_1}{\partial y} \, dx \, dy$$

$$= \frac{1}{4\Delta} [\beta_2\beta_1 + a\beta_2\gamma_1 - a\gamma_2\beta_1 + \gamma_1\gamma_2] .$$

In the same way,

$$(K_{13})_e = \frac{1}{4\Delta}\left[\beta_1\beta_3 + \gamma_1\gamma_3 + a\gamma_1\beta_3 - a\gamma_3\beta_1\right] ,$$

$$(K_{21})_e = \frac{1}{4\Delta}\left[\beta_2\beta_1 - a\beta_2\gamma_1 + a\gamma_2\beta_1 + \gamma_1\gamma_2\right] ,$$

$$(K_{22})_e = \frac{1}{4\Delta}\left[\beta_2^2 + \gamma_2^2\right] ,$$

$$(K_{23})_e = \frac{1}{4\Delta}\left[\beta_2\beta_3 + \gamma_2\gamma_3 + a\gamma_2\beta_3 - a\gamma_3\beta_2\right] ,$$

$$(K_{31})_e = \frac{1}{4\Delta}\left[\beta_1\beta_3 + \gamma_1\gamma_3 - a\gamma_1\beta_3 + a\gamma_3\beta_1\right] ,$$

$$(K_{32})_e = \frac{1}{4\Delta}\left[\beta_2\beta_3 + \gamma_2\gamma_3 - a\gamma_2\beta_3 + a\gamma_3\beta_2\right] ,$$

$$(K_{33})_e = \frac{1}{4\Delta}\left[\beta_3^2 + \gamma_3^2\right] .$$

5.2.2 Assemblage of the local matrices

The remaining step is to assemble matrices from each element into one global matrix. If the element's global node number is i, j and k respectively then

$$\begin{aligned} K_{ii} &= (K_{11})_e , \\ K_{ij} &= (K_{12})_e , \\ K_{ik} &= (K_{13})_e , \\ K_{ji} &= (K_{21})_e , \\ K_{jj} &= (K_{22})_e , \\ K_{jk} &= (K_{23})_e , \end{aligned} \qquad (5.31)$$

$$K_{ki} = (K_{31})_e ,$$

$$K_{kj} = (K_{32})_e ,$$

$$K_{kk} = (K_{33})_e ,$$

$$Q_i = (Q_1)_e ,$$

$$Q_j = (Q_2)_e ,$$

$$Q_k = (Q_3)_e .$$

The contribution from each element in the process of the assemblage is illustrated in Fig. 5.6 [47].

The constant boundary condition (5.9) is forced in the procedure to solve the global matrix equations.

6. ANALYSIS OF THE VERTICAL HALL EFFECT DEVICES

6.1 Introduction.

Conventional Hall effect devices have been in the form of a plate situated in the plane of the chip surface, hence becoming sensitive to a magnetic field perpendicular to the active chip surface. However some applications need devices sensitive to a magnetic field parallel to the device surface. For those applications, the VHD recently proposed by Popović [48] seems better than the magnetotransistor or the magnetodiode in the sense of temperature dependence, linearity and noise dependence [35].

Figure 6.1 (a) shows the Hall effect device which is a thin plate of a semiconductor with four ohmic contacts, one on each edge of the plate. However it is impossible to integrate such a vertical plate because only one contact is available on the surface. If we can deform the Hall plate until both sensor contacts, S1 and S2, come to the plane where the current contact C1 is, then we have three contacts on the chip surface (Fig. 6.1 (b) and (c)). Even then it is not easy to integrate a perfectly semicircular device so a rough approximation of the semicircular device is taken (Fig. 6.2), which pratically has the same performance [48].

Recently Huiser and Baltes [35,36] simulated the VHD using Green's functions. In this thesis, the numerical simulation of the VHD using a finite element method is presented.

6.2 Equations

A two dimensional geometry of the device is shown in Fig. 6.2. The x-axis is parallel to the surface of the chip and perpendicular to the magnetic induction $\vec{B}$. The y-axis is perpendicular to the surface of the chip.

Neglecting diffusion currents, the electron current density equation in the presence of a not too large field ($\mu^* B \leq 0.3$) now becomes [17]

$$\vec{J}_n = q\mu_n n\vec{E} - \mu_n^* (\vec{J}_n \times \vec{B}) \quad . \tag{6.1}$$

Calculating the divergence of (6.1) and using the relation curl grad V = 0, we arrive at [35]

$$\text{div } J_n = - [\ q\mu_n n \ / \ (1 + (\mu_n{}^* B)^2)\ ]\ \text{div grad } V \quad , \tag{6.2}$$

where μ_n and $\mu_n{}^* B$ are assumed to be x and y independent. Since no injection of electrons or holes takes place in the bulk, div $\vec{J}_n = 0$.

Hence, the resulting equation becomes Laplace's equation,

$$\text{div grad } V = 0 \ . \tag{6.3}$$

The current normal to the boundary where there are no electrodes, must vanish. Using (3.17) and (3.18) this implies that [49,50]

$$[\ \frac{\partial V}{\partial x} - \mu_n{}^* B \frac{\partial V}{\partial y} \ , \ \frac{\partial V}{\partial y} + \mu_n{}^* B \frac{\partial V}{\partial x}\ ] \cdot \vec{n} = 0 \quad , \tag{6.4}$$

where $\vec{n}$ is the unit normal vector to the boundary.

When we compare equations (6.3) and (6.4) with equations (5.8) to (5.10) we notice that the two sets of equation are the same if f = 0, a = $\mu_n{}^* B$ and ϕ = V are assumed. Hence, the discretized equations from Chapter 5 are used to simulate the VHD under the above assumptions.

6.3 Solution Procedure

The grid for the simulation of the VHD is generated (Fig. 6.3). with a dense network of nodes in the surface region of the device because of reasons explained in Chapter 4. The number of nodes is 682, which is the number of unknown variables, and the number of elements is 1128.

The K matrix which is in LHS of (5.29) is then generated. The components of the Q matrix which is RHS of (5.29) are zero because in our equation f = 0. These element K

matrices are assembled globally according to (5.31). Now the global matrix is solved by an adaptive SOR method. Since the size of the global matrix is large, direct methods such as Gaussian Ellimination are not efficient. The adaptive SOR is approximately 5 times faster than normal SOR in our problem.

To operate the VHD in the current source mode, a linear combination of operating voltages is used in determining the voltage distribution in the device. $V_2(x,y)$ represents the voltage of the VHD when we supply 1 volt to contact 2 and 0 volts to the rest of the contacts. The same interpretation may be applied to the $V_3(x,y)$ and $V_4(x,y)$.

The resulting V(x,y) reads

$$V(x,y) = k_2 V_2(x,y) + V_3(x,y) + k_4 V_4(x,y) \tag{6.5}$$

Integrating V(x,y) over the sensor contact 2(a-b) and 4(c-d) equation (6.5) reads,

$$\int_a^b \frac{\partial V(x,y)}{\partial y}\,dx = k_2 \int_a^b \frac{\partial V_2(x,y)}{\partial y}\,dx + \int_a^b \frac{\partial V_3(x,y)}{\partial y}\,dx + k_4 \int_a^b \frac{\partial V_4(x,y)}{\partial y}\,dx \tag{6.6}$$

$$\int_c^d \frac{\partial V(x,y)}{\partial y}\,dx = k_2 \int_c^d \frac{\partial V_2(x,y)}{\partial y}\,dx + \int_c^d \frac{\partial V_3(x,y)}{\partial y}\,dx + k_4 \int_c^d \frac{\partial V_4(x,y)}{\partial y}\,dx$$

By controlling the LHS of equation (6.6) the VHD may be used in the current source mode and a Hall voltage, which is (k_2-k_4), is determined. To obtain V(x,y) the global matrix has to be solved three times, each time with different boundary conditions. To see if the true solution was actually obtained for the given equation, the conservation of the terminal

currents is checked.

From the simulation results associated with different sensor probe positions, the optimum position which gives the highest Hall voltage may be determined.

The flow chart for the solution procedure is shown in Fig. 6.9 (a), (b) and (c).

6.4 Results and Discussion

In all our calculations we assume that the free carriers are electrons and that the doping concentration is $N_D = 10^{15}$ cm^{-3}. We also assume the electron mobility to be 1350 cm^2/Vs. For n-type silicon, r_n is about 1.15 at room temperature for low donor concentration [4,16]. Thus the electron Hall mobility is assumed as 1.15 times the electron mobility. The device geometry is 120 μm wide, 70 μm deep and 12 μm thick. Each contact has 10 μm length.

First, we tried to find the optimum position of the sensor probes 2 and 4. The Hall voltage at 7 different positions (Table 6.1) was compared in order to determine an optimum probe position. Figure 6.4 shows the Hall voltage at these different positions for μ*B = 0.01, 0.1 and 0.2. Current conservation should be noted from Table 6.2 which shows terminal voltages and currents for μ*B = 0.2. Conventionally the optimum position has been believed to be located at the middle of the two fixed contacts (1,3 and 3,5, i.e. position 4). However, our results indicate the Hall voltage is maximum at the vicinity of position 6 regardless of magnetic induction B. The results shown hereafter are for the probe fixed at position 6.

Figure 6.5 shows the potential distribution in the VHD for μ*B = 0.2 with 1 volt applied at electrode 3 while electrode 1 and 5 being grounded. The electrodes 2 and 4 are the probes where the Hall voltage developes in the presence of a magnetic field, when the device operates in the current source mode (zero current). The ideal Hall voltage is 0.1474 V (see Chapter 2), the actual Hall voltage is 0.0889 V and the efficiency is 60.3 %. The sensitivity defined in terms of V/AT is 484 (V/AT) which is well above that obtained from Popović's measurement [48]. The value of efficiency from Huiser and Baltes [35,36] shows a slightly higher value ranging from 62.5 % to 64.8 % if devices having 5 contacts are considered. The

value of sensitivity from Huiser and Baltes is also slightly higher than our value by 15 to 40 (V/AT).

The equipotential lines shown in Fig. 6.6 are for the VHD for $\mu^*B = 0.2$. We can notice an asymmetry in the picture, caused by the magnetic field. The equipotential and current lines from Huiser and Baltes are shown in Fig. 6.7. Figure 6.8 shows the Hall voltage vs. magnetic induction B. The Hall voltage is a linear for μ^*B less than 0.2.

7. ANALYSIS OF THE MAGFET

The MAGFET is a MAGnetic-field sensitive Field Effect Transistor. There are two types of MAGFET. The first kind is a single drain structure, based predominantly on the Hall effect, with sensor probes on both sides of the gate so that a Hall voltage is detected. The second kind is a multiple drain structure whose sensitivity to the magnetic field is based on the carrier deflection effect. The sensitivity is a function of the current imbalance between the drains. We will consider a double drain and a triple drain device structure.

The MAGFET, just as the ordinary MOSFET has several advantages ;

4. it can be easily integrated with other electronic circuits on a single chip and fabricated by the standard CMOS process,
5. low power consumption.

However, it has the following disadvantages ;

1. since the action takes place near the surface (inversion layer) there are noise problems and lower sensitivities because of the low mobility,
2. offset problems due to nonideal device geometries.

Consider an N-channel MOSFET operated in the linear region as shown in Fig. 7.1 [16]. The magnetic-field, $\vec{B}$ is assumed to be perpendicular to the channel, i.e. $\vec{B} = (0,0,B_z)$; thus the Lorentz force acts in the plane of the channel (x-y plane). As in the case of zero magnetic-field, we assume no current flow in the z-direction [5].

7.1 Assumptions

In the following, the assumptions made in the analysis of the MAGFET are discussed [5,16] :

a. the gate structure corresponds to an ideal MOS diode so that any interface traps, fixed oxide charges or work function differences can be neglected. In a nonideal gate, the effect of fixed charges and difference in work functions cause a voltage shift corresponding to the flatband voltage, V_{FB}. This in turn causes a change in the threshold voltage, V_T.

b. the doping in the channel is uniform.

c. the drift and Hall mobilities in the inversion layer are constants for given gate and drain bias voltages. This is done so by neglecting $(\partial\mu_n/\partial z)$, $(\partial\mu_n^*/\partial z)$ across the channel and assuming no lateral dependence of μ_n, μ_n^* (for MOSFET's operating in the linear region under low drain bias voltages, $E_z(x,y)$ can be assumed to be constant).

d. the reverse leakage current is negligibly small.

e. the transversal field E_z is much greater than the longitudinal field E_y in the channel. This corresponds to the so-called Gradual Channel Approximation, which assumes that the charges in the system are governed by the vertical field only.

f. within the inversion layer, the hole current density can be neglected since $\vec{J}_p << \vec{J}_n$ for $V_D \geq 0$ and $V_B \leq 0$. The net recombination rate can be assumed zero in the channel. This yields the drain current, I_D, to be an electron current of vanishing divergence.

7.2 Basic Equations

With the assumptions in the previous section, the charge in the inversion layer, if we consider only drift current, is given by [5,16,51],

$$Q_n(x,y) = C_{ox}\,[\,V(x,y) + 2\psi_B - V_G\,] + \sqrt{2\varepsilon_{Si}qN_A\,[\,V(x,y) + 2\psi_B\,]}\,, \tag{7.1}$$

where $V(x,y)$ denotes the reverse bias voltage between a point (x,y) in the channel and the source electrode which is at zero volts. C_{ox} is the gate oxide capacitance per unit area ;

$$C_{ox} = \frac{\varepsilon_{SiO_2}}{t_{ox}}\,, \tag{7.2}$$

where t_{ox} is the oxide thickness. ψ_B is the potential difference between the bulk Fermi level (E_{Fp}) and the intrinsic Fermi level (E_i). ψ_B is given by

$$\psi_B = \frac{kT}{q}\log_e\left(\frac{N_A}{n_i}\right)\,, \tag{7.3}$$

where kT/q = 25.7 mV at room temperature, N_A is the bulk doping density and n_i is the intrinsic concentration.

From (3.17) and (3.18), the two dimensional current density equations, using $\vec{E} = -\vec{\nabla}V$ and $D_n = \mu_n kT/q$, reads

$$J_{nx}(x,y) = \frac{\mu_n}{1 + (\mu_n^* B_z)^2} \left\{ -\frac{\partial V(x,y)}{\partial x} + \frac{kT}{q}\frac{\partial}{\partial x} \right. \tag{7.4}$$

$$\left. -\mu_n^* B_z \left[-\frac{\partial V(x,y)}{\partial y} + \frac{kT}{q}\frac{\partial}{\partial y} \right] \right\} qn(x,y) ,$$

$$J_{ny}(x,y) = \frac{\mu_n}{1 + (\mu_n^* B_z)^2} \left\{ -\frac{\partial V(x,y)}{\partial y} + \frac{kT}{q}\frac{\partial}{\partial y} \right. \tag{7.5}$$

$$\left. +\mu_n^* B_z \left[-\frac{\partial V(x,y)}{\partial x} + \frac{kT}{q}\frac{\partial}{\partial x} \right] \right\} qn(x,y) .$$

By integrating the electron current density equations over the channel depth t, we obtain line current density equations for the MAGFET. This is carried out using assumptions discussed previously (in particular assumption (c)) and further neglecting the dependence of $(\partial V/\partial x)$, $(\partial V/\partial y)$ on z and neglecting diffusion current terms [5];

$$\hat{J}_{nx}(x,y) = \int_0^t J_{nx}(x,y,z)\, dz \tag{7.6}$$

$$= \frac{-1}{1 + (\mu_n^* B_z)^2} \left[\mu_n \frac{\partial V}{\partial x} - \mu_n^* B_z \left(\mu_n \frac{\partial V}{\partial y}\right) \right] Q_n(x,y) ,$$

$$\hat{J}_{ny}(x,y) = \int_0^t J_{ny}(x,y,z)\, dz \tag{7.7}$$

$$= \frac{-1}{1 + (\mu_n^* B_z)^2} \left[\mu_n \frac{\partial V}{\partial y} + \mu_n^* B_z \left(\mu_n \frac{\partial V}{\partial x} \right) \right] Q_n(x,y) \quad ,$$

$$Q_n(x,y) = q \int_0^t n(x,y,z)\, dz \quad . \tag{7.8}$$

The integration is carried out from the semiconductor-oxide interface ($z = 0$) to the point ($z = t$) at which E_i intersects E_{Fn} (the quasi-Fermi level for electrons).

Integrating (7.8) from the source (0 V) to a potential V, at any arbitrary point (x,y) in the channel [5], viz.

$$\int_0^V Q_n(V')\, dV' = \tilde{Q}_n(V) \quad , \tag{7.9}$$

$$\frac{\partial \tilde{Q}_n(V)}{\partial x} = \frac{\partial \tilde{Q}_n(V)}{\partial V} \frac{\partial V}{\partial x} = Q_n(V) \frac{\partial V}{\partial x} \quad , \tag{7.10}$$

$$\frac{\partial \tilde{Q}_n(V)}{\partial y} = Q_n(V) \frac{\partial V}{\partial y} \quad . \tag{7.11}$$

Substituting (7.10) and (7.11) into (7.6) and (7.7) and introducing a new function

$$\phi = -\tilde{Q}_n(V) \quad , \tag{7.12}$$

we can simplify (7.6), (7.7) as follows :

$$\hat{J}_{nx}(x,y) = \frac{\mu_n}{1 + (\mu_n^* B_z)^2} \left(\frac{\partial \phi}{\partial x} - \mu_n^* B_z \frac{\partial \phi}{\partial y} \right) \quad , \tag{7.13}$$

$$\hat{J}_{ny}(x,y) = \frac{\mu_n}{1 + (\mu_n^* B_z)^2} \left(\frac{\partial \phi}{\partial y} + \mu_n^* B_z \frac{\partial \phi}{\partial x}\right) \quad (7.14)$$

The electron continuity eqn.(3.25), using assumption (f) becomes

$$\frac{\partial}{\partial x} \hat{J}_{nx}(x,y) + \frac{\partial}{\partial y} \hat{J}_{ny}(x,y) = 0 \quad (7.15)$$

Hence from (7.13), (7.14) and (7.15) we have

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0 \quad (7.16)$$

where [5,32]

$$\phi = -\frac{1}{2} V^2 C_{ox} + C_{ox} (V_G - 2\phi_B)V \quad (7.17)$$

$$- \frac{1}{2\varepsilon_{Si} q N_A} (2\varepsilon_{Si} q N_A V + 4\varepsilon_{Si} q N_A \phi_B)^{1.5}$$

7.3 Boundary Conditions

At the floating boundary ③, in Fig. 7.3 and Fig. 7.4, $J_{nx} = 0$.

Hence,

$$\frac{\partial \phi}{\partial y} + \mu_n^* B_z \frac{\partial \phi}{\partial x} = 0 \quad . \tag{7.18}$$

At the floating boundaries ① and ② in Fig. 7.3 and Fig. 7.4, $J_{ny} = 0$.

Hence,

$$\frac{\partial \phi}{\partial x} - \mu_n^* B_z \frac{\partial \phi}{\partial y} = 0 \quad . \tag{7.19}$$

For the contact boundary condition, the ohmic contact which has infinite recombination velocity and space charge neutrality is assumed,

$$\phi_D = \phi|_{V=V_D} \quad , \tag{7.20}$$

$$\phi_S = \phi|_{V=V_S} \quad . \tag{7.21}$$

7.4 Numerical Method

The equation for the MAGFET (7.16) is Laplace's equation with boundary condition (7.18), (7.19), (7.20) and (7.21). When we compare (7.16), and (7.18) through (7.21) with the equations from (5.8) to (5.10) we notice that the two sets of equations are the same if $f = 0$, $a = \mu^* B_z$ and $V = \phi$. Thus we use the discretized equations from Chapter 5 to simulate the MAGFET under the assumption that $f = 0$, $a = \mu^* B_z$ and $V = \phi$. However we have to solve (7.17) to get the value of V(x,y) by using the Newton-Raphson algorithm [52].

7.5 Solution Procedure

The grid for the simulation of the MAGFET is generated as explained in Chapter 4 (Fig. 7.2). The nodes at the drain regions are dense because of singular points. The number of nodes is 633 and the number of elements is 1060. Having generated the grid, the K matrix which is in LHS of (5.29) is formed. The element K matrices are assembled globally according to (5.31). The global matrix is solved by an adaptive SOR method.

As for the initial guess the variable ϕ is provided with linear values, i.e.

$$\phi^{o}(I) = [\ (\phi_D - \phi_S)Y(I) \ / \ (YMAX - YMIN)\ ] + \phi_S \ , \qquad (7.22)$$

where I is the node number and Y(I) is the y coordinate of the node I. YMAX is the maximum y coordinate and YMIN is the minimum y coordinate.

For the Newton-Raphson iterations, V is provided with linear guess values, i.e.

$$V^{o}(I) = [\ (V_D - V_S)Y(I) \ / \ (YMAX - YMIN)\ ] + V_S \ , \qquad (7.23)$$

where V_D is the drain voltage and V_S is the source voltage.

The flowchart for the solution procedure is shown in Fig. 7.11 (a) and (b).

7.6 Results and Discussion

In all our calculations we assume an N-channel MAGFET with substrate doping, $N_A = 10^{15}$ cm^{-3}, and oxide thickness, t_{ox} = 100 nm. The magnetic field is perpendicular to the device plane. The device is in the linear region of operation and threshold voltage, V = 1.1 V. We also assume 750 cm^2/Vs for the electron surface mobility and 1.15 for r_n. The double drain MAGFET and triple drain MAGFET were simulated and their sensitivities were compared. The sensitivity defined in terms of the relative current imbalance between the two outer drains (see Fig. 7.3-7.4) is defined as [5]

$$\delta = (I_{D1} - I_{D2}) \ / \ (I_{D1} + I_{D2}) \ . \qquad (7.24)$$

The device geometry for the split drain structure is 100 μm wide and 100 μm long. This geometry seems to be favourable in view of the sensitivity, signal to noise ratio and design rules [5]. The device is operated in the linear region, $V_{D1S} = V_{D2S} = V_{D3S} = 1$ V, $V_{GS} = 3$ V and $V_S = 0$ V. The top view of the double drain MAGFET is given in Fig. 7.3. The length of the drain contacts is 20 μm each and the separation distance d is 60 μm. Figure 7.4 shows the top view of the triple drain MAGFET. The length of the drain contacts is 20 μm each and the seperation distance is 20 μm each.

Figure 7.5 shows equipotential lines in the channel of the double drain MAGFET for $\mu^*B = 0.1$. We notice an asymmetry along the y axis caused by the magnetic field. This results in the potential lines having a skew of the Hall angle at the floating boundaries. Equipotential lines for $\mu^*B = 0.2$ are shown in Fig. 7.6. The Hall angle is smaller than that in Fig. 7.5. Figure 7.7 illustrates the potential distribution of the double drain MAGFET for $\mu^*B = 0.2$. The equipotential lines and potential distributions of the triple drain MAGFET for $\mu^*B = 0.2$ are shown in Fig. 7.8 and Fig. 7.9 respectively.

From the sensitivity curve (Fig. 7.10), we can conclude that the sensitivity is almost linear in the range from -2 tesla to 2 tesla. In view of the sensitivity values, the triple drain MAGFET is superior to the double drain MAGFET.

8. CONCLUSION AND OUTLOOK

To develop new semiconductor devices, the understanding of device behaviour under general geometries and operating conditions is inevitable. In the realm of analytical analysis only limiting cases can be understood. The analysis of semiconductor devices under the influence of a magnetic field becomes even more difficult using analytical methods because the magnetic field introduces asymmetry into the current density equations and boundary conditions.

In this thesis, the numerical modeling of semiconductor devices such as the VHD and MAGFET is performed. Among the several numerical methods available to solve partial differential equations, the finite element method is used.

To support the finite element algorithm efficiently, an adaptive grid generation method is developed. Adaptive grid generation is one of the most desired but difficult parts of the finite element method [18]. With the aid of a proper grid, the VHD and triple drain MAGFET can be simulated. The initial grid generation procedure for the arbitrary device geometries will be modified to provide more generality and convenience to the user by incorporating with pointing devices such as mouse or optical pen.

For the simulation of the VHD and MAGFET, only the electron continuity equation is used. In order to simulate more complex devices such as the Lateral Magnetotransistor (LMT) and the magnetodiode three equations are necessary, viz., the continuity equations for electrons and holes and Poisson equation. The program is presently being extended to incorporate these three equations.

Adding a powerful graphic and interactive user interface for device specification and analysis of results becomes a necessity, in particular when it should be connected any other VLSI design package. With the extensions stated above, the range of semiconductor devices which this program can handle will be increased.

The results from numerical device modeling need not perfectly agree with those obtained from measurements on real devices. This is due to several assumptions made in

modeling, process errors such as geometry offset and/or piezoresistive effect of Si. Furthermore, in the fabricated device there are several uncertain factors such as noise, temperature dependence and frequency dependence. Incorporating these effects into a simulation program is highly likely in the forseeable future because it will enhance the accuracy of device modeling results. However this trend will be governed on the basis of performance/effort ratio. If all of the program modules are networked under A.I. (Artificial Intelligence) environment, then the man-month required for the device modeling can be significantly reduced.

Position	Y coordinate (μm)			
	a	b	c	d
1	20	30	90	100
2	22.5	32.5	87.5	97.5
3	25	35	85	95
4	27.5	37.5	82.5	92.5
5	30	40	80	90
6	32.5	42.5	77.5	87.5
7	35	45	75	85

Table 6.1 Position of the electrodes

Electrode	1	2	3	4	5
V (V)	0	0.4783	1	0.5672	0
I (A/cm)	-0.0757	0	0.1592	0	-0.0835

Table 6.2 Voltages and currents of the electrodes at position 6

(μ*B = 0.2)

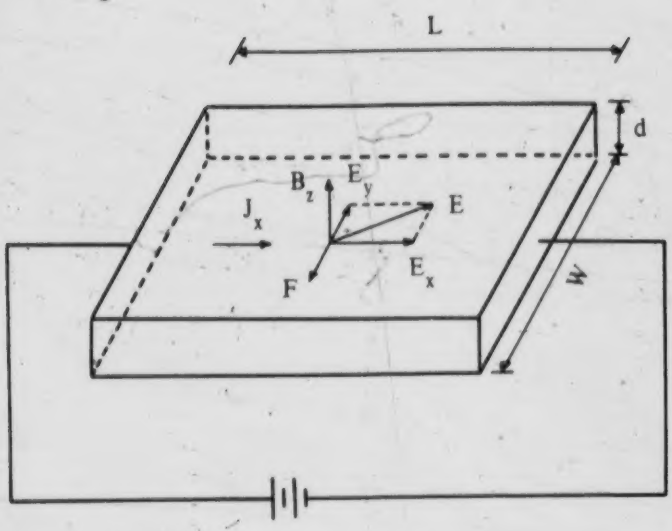


Fig. 2.1 (a) The Hall effect in a p-type semiconductor

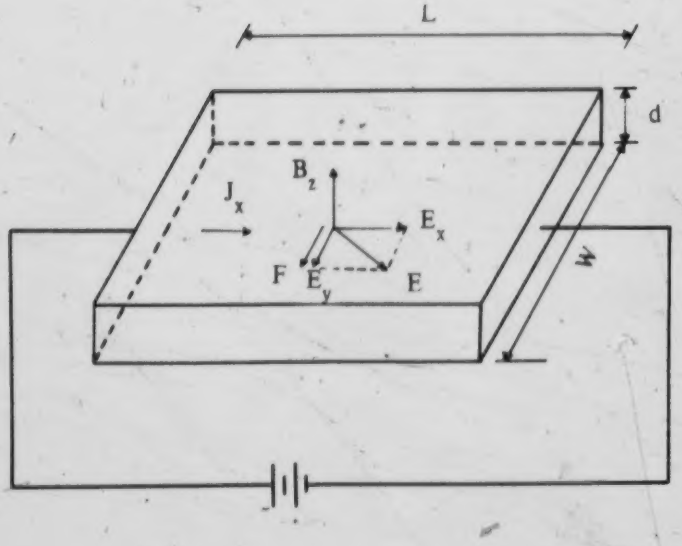


Fig. 2.1 (b) The Hall effect in a n-type semiconductor

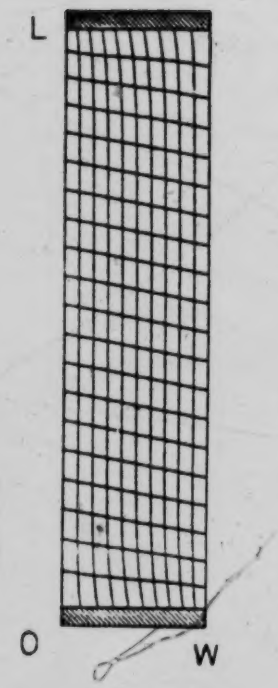


Fig. 2.2 Equipotential and current lines (L = 4W) [5]

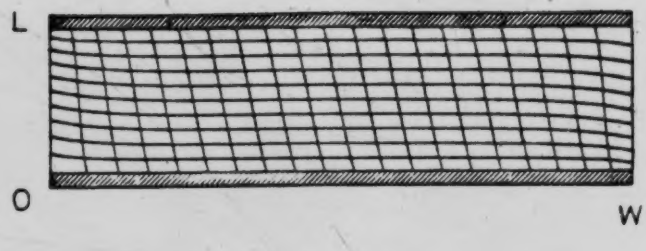


Fig. 2.3 Equipotential and current lines (W = 4L) [5]

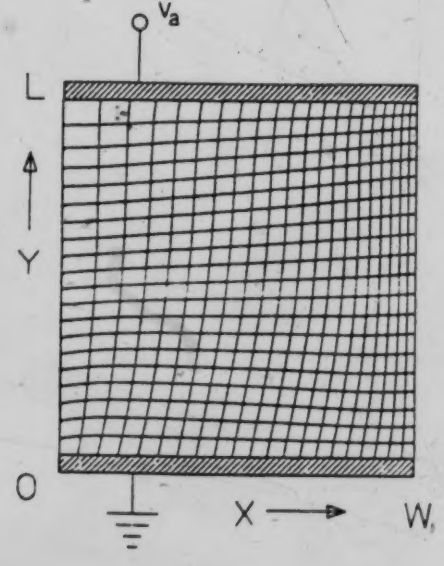


Fig. 2.4 Equipotential and current lines of intrinsic square device [5]

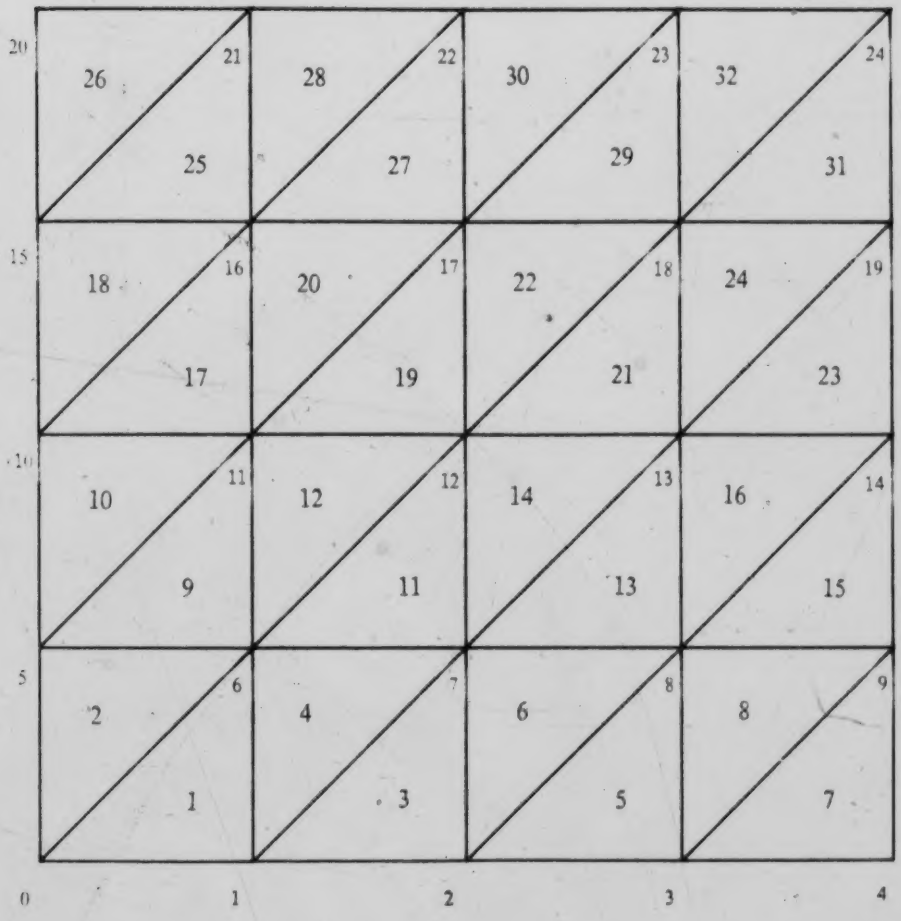


Fig. 4.1 The numbering scheme for the initial grid

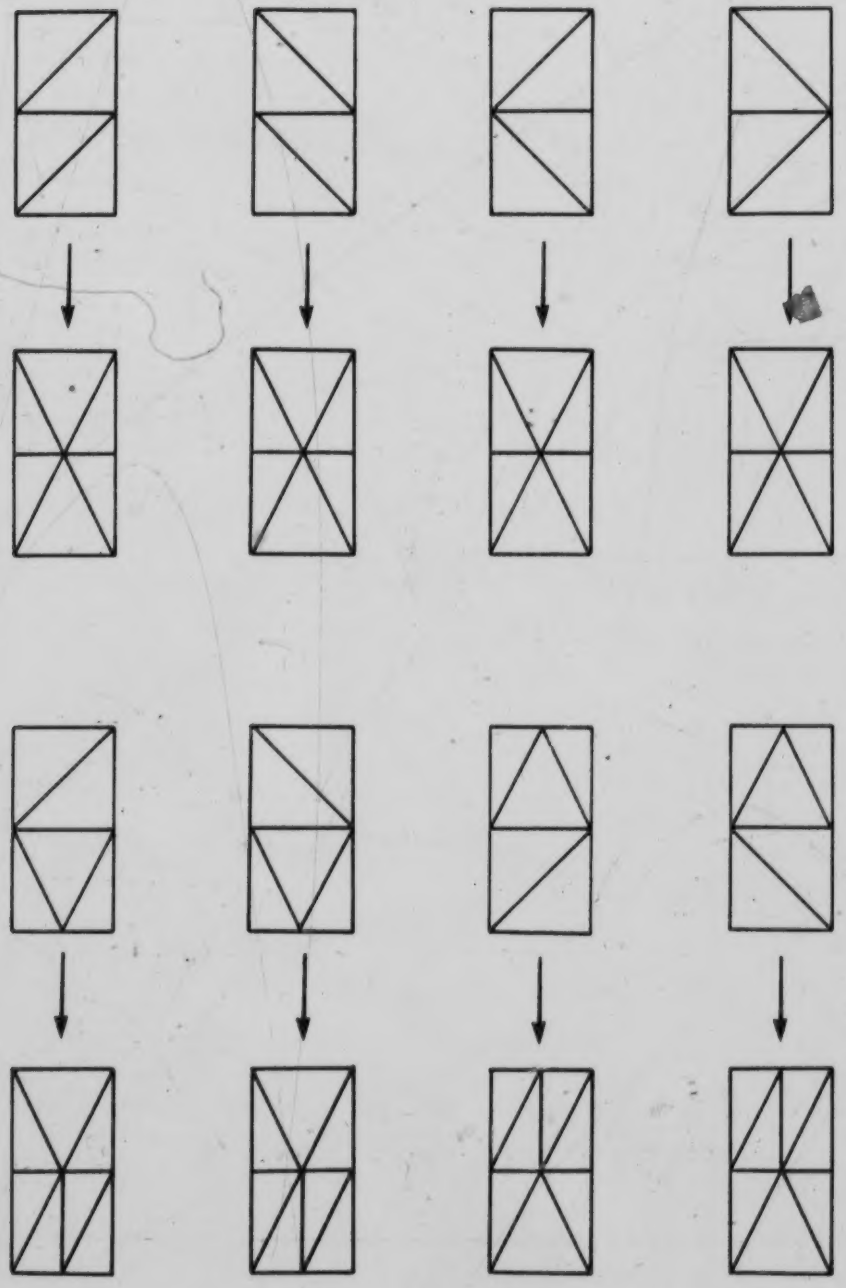


Fig. 4.2 (a) Refinable element shapes for horizontal refinement pass

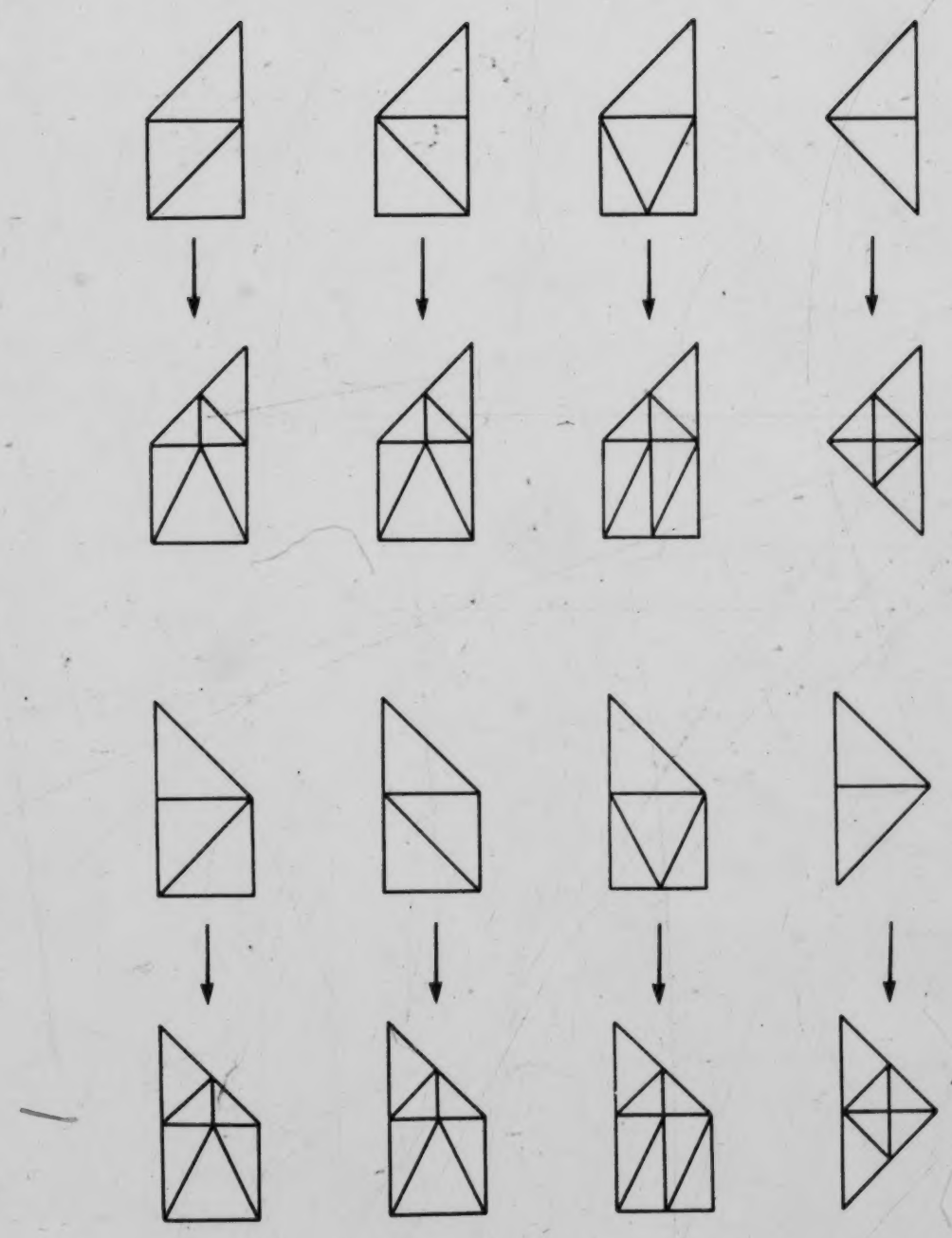


Fig. 4.2 (b) Refinable element shapes for horizontal refinement pass

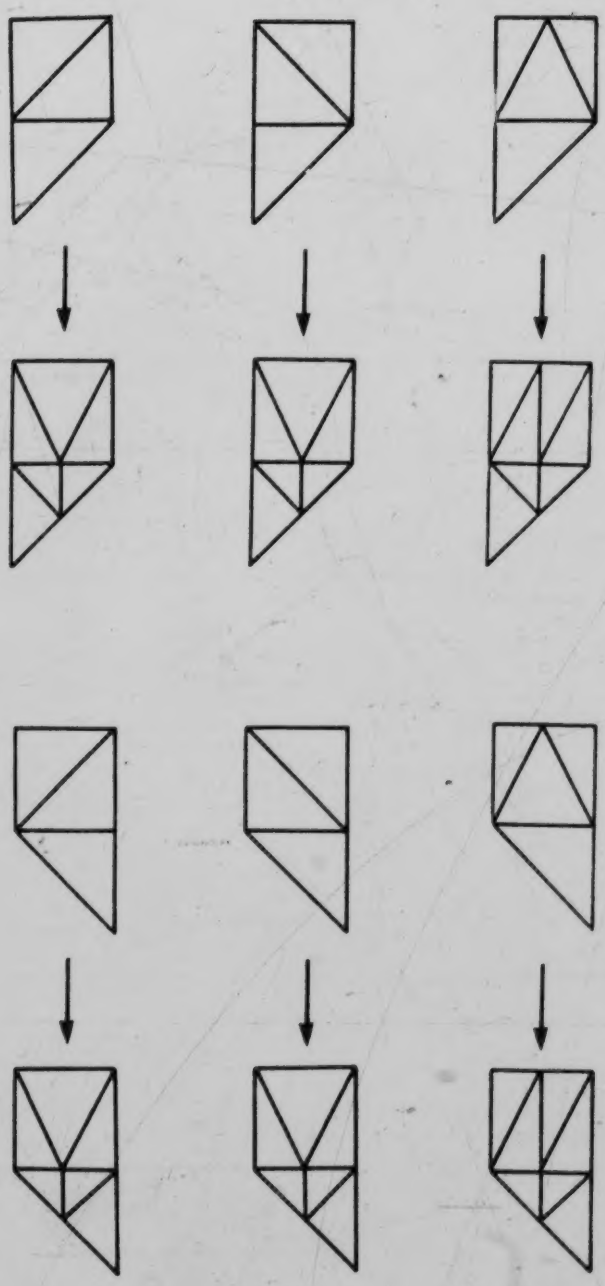


Fig. 4.2 (c) Refinable element shapes for horizontal refinement pass

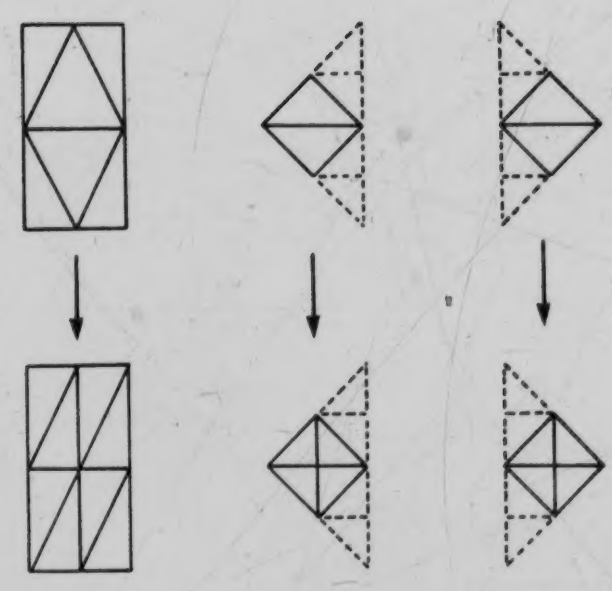


Fig. 4.2 (d) Refinable element shapes for horizontal refinement pass

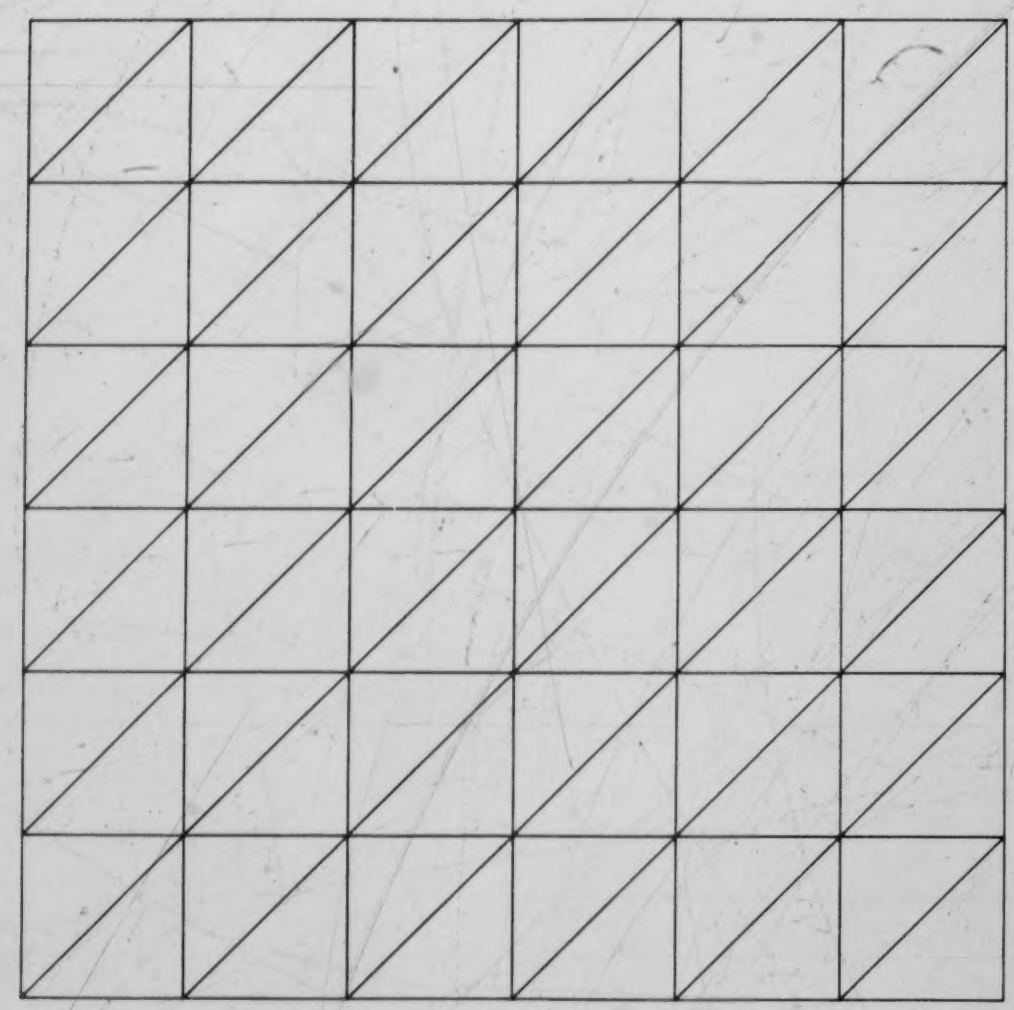


Fig. 4.3 A grid for a rectangular region

(a) An initial grid

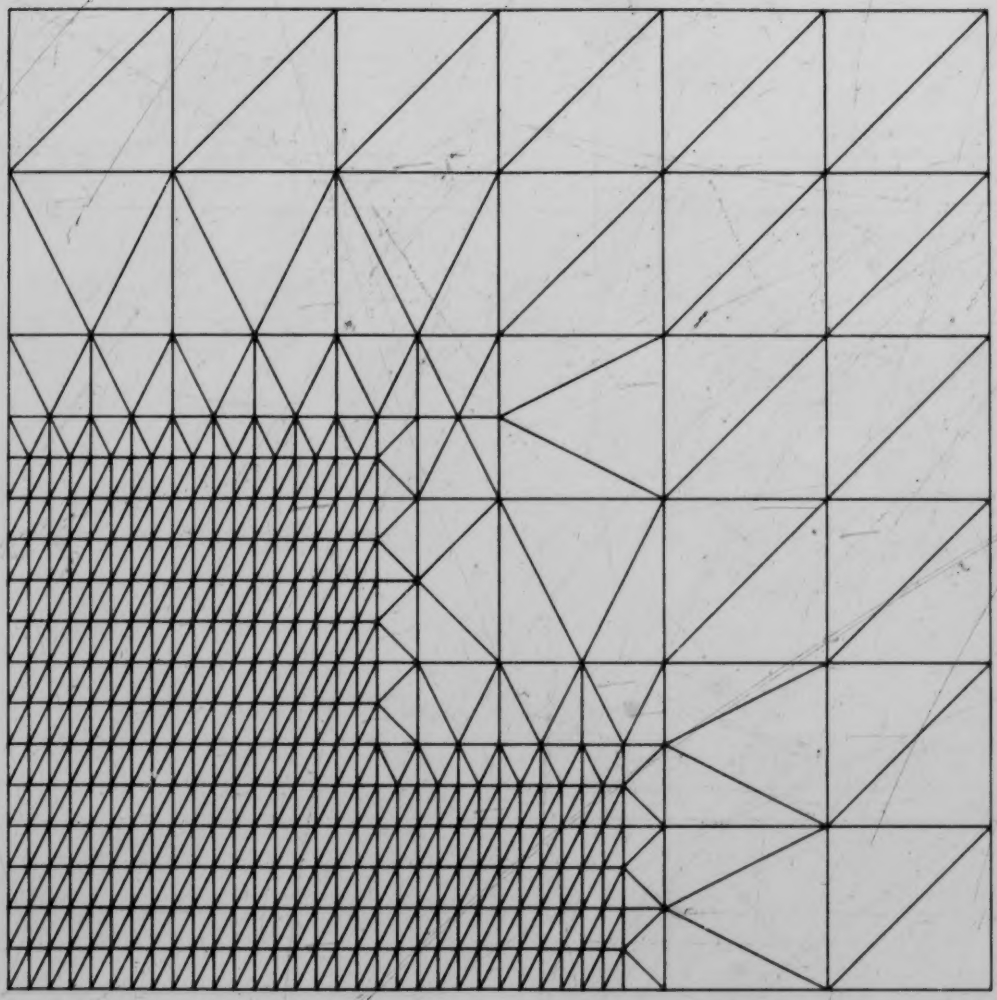
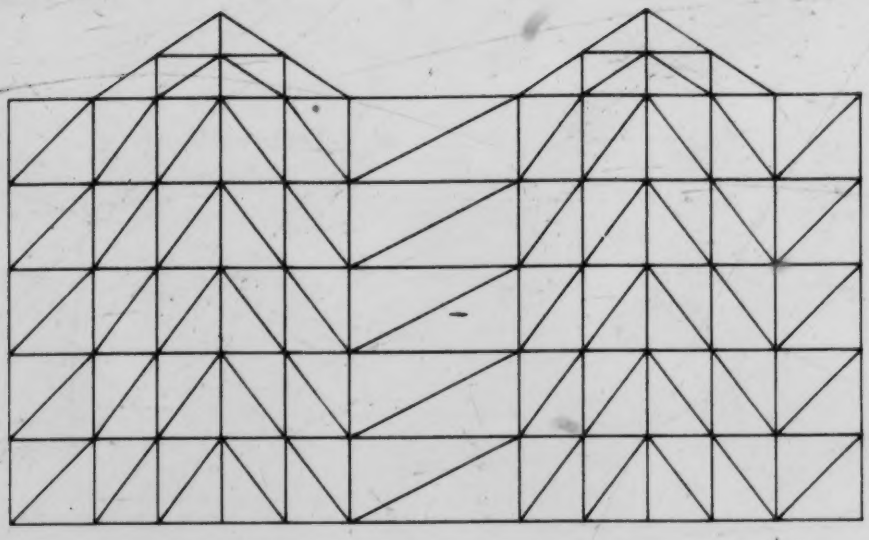
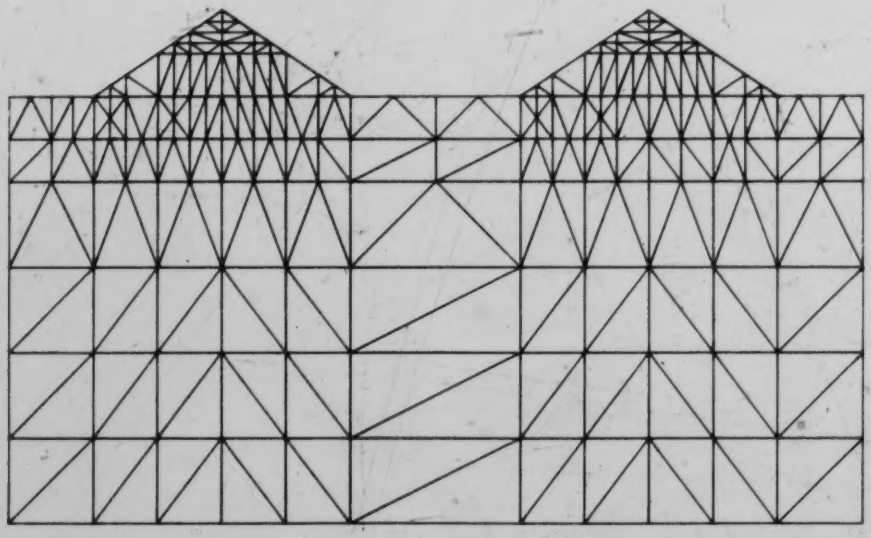


Fig. 4.3 A grid for a rectangular region

(b) A final grid with 400 nodes



(a)

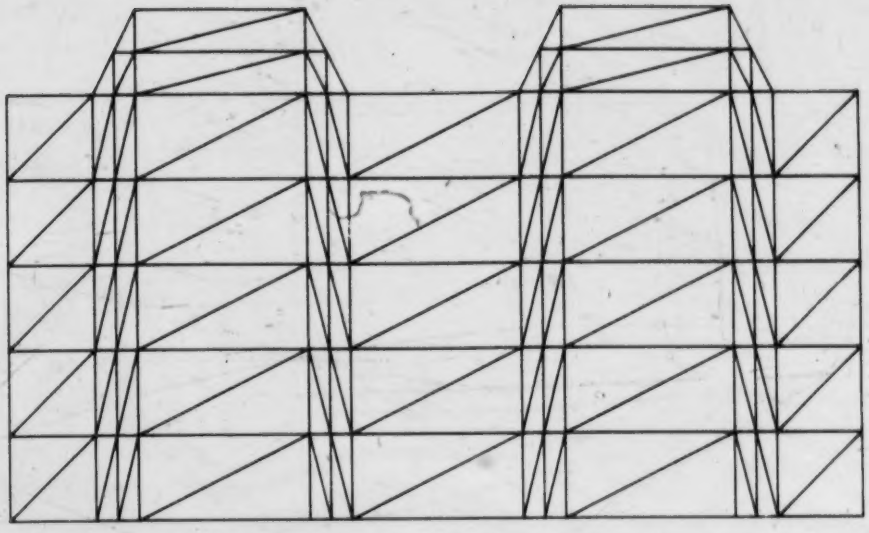


(b)

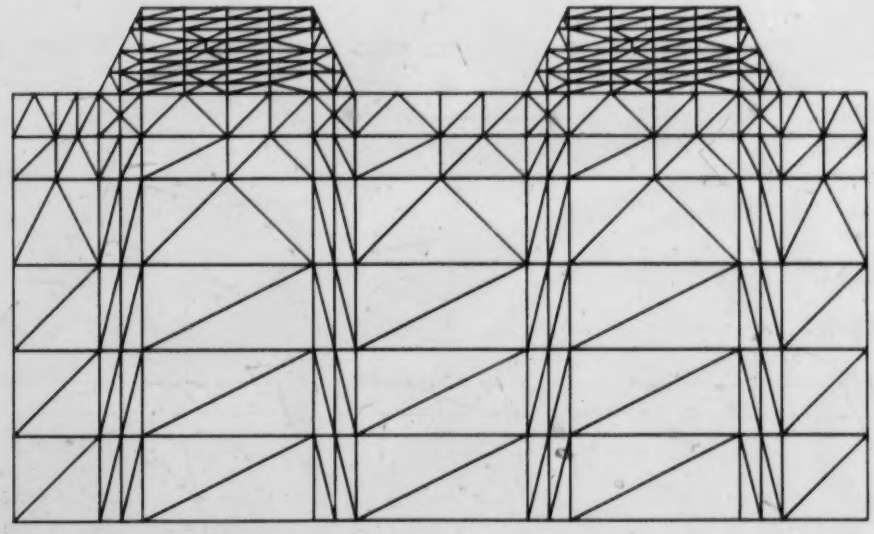
Fig. 4.4 The first example of a grid for a region having flat bottom.

(a) An initial grid.

(b) A final grid with 200 nodes.



(a)



(b)

Fig. 4.5 The second example of a grid for a region having flat bottom.

(a) An initial grid.

(b) A final grid with 230 nodes.

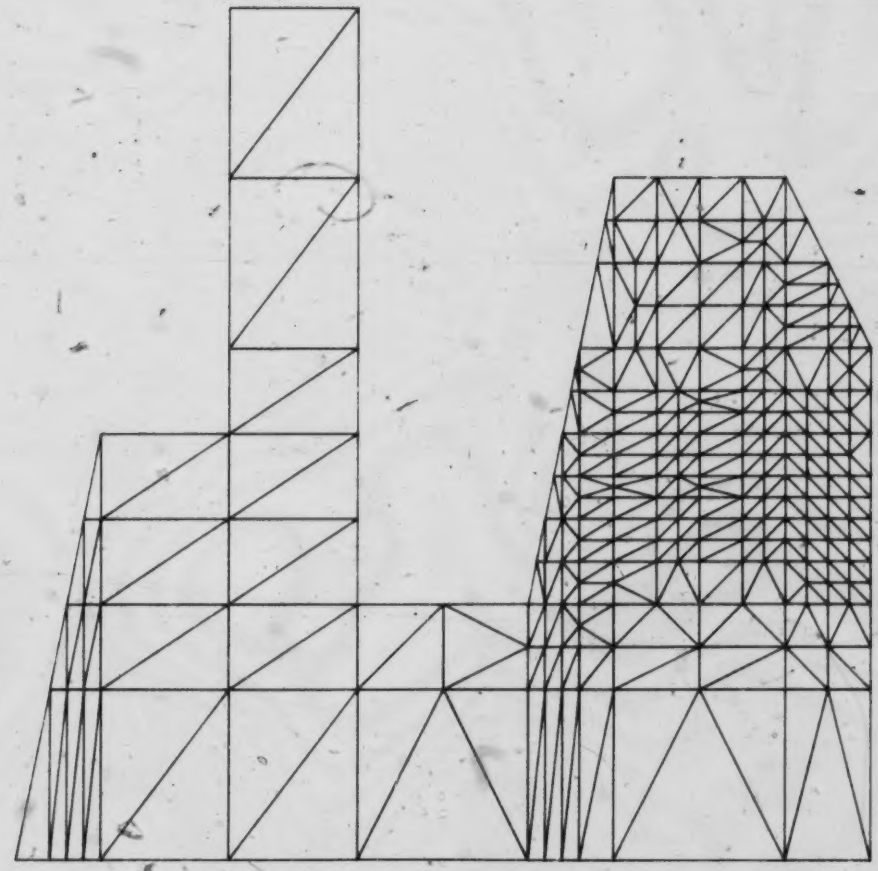


Fig. 4.6 The third example of a grid for a region having flat bottom with 250 nodes.

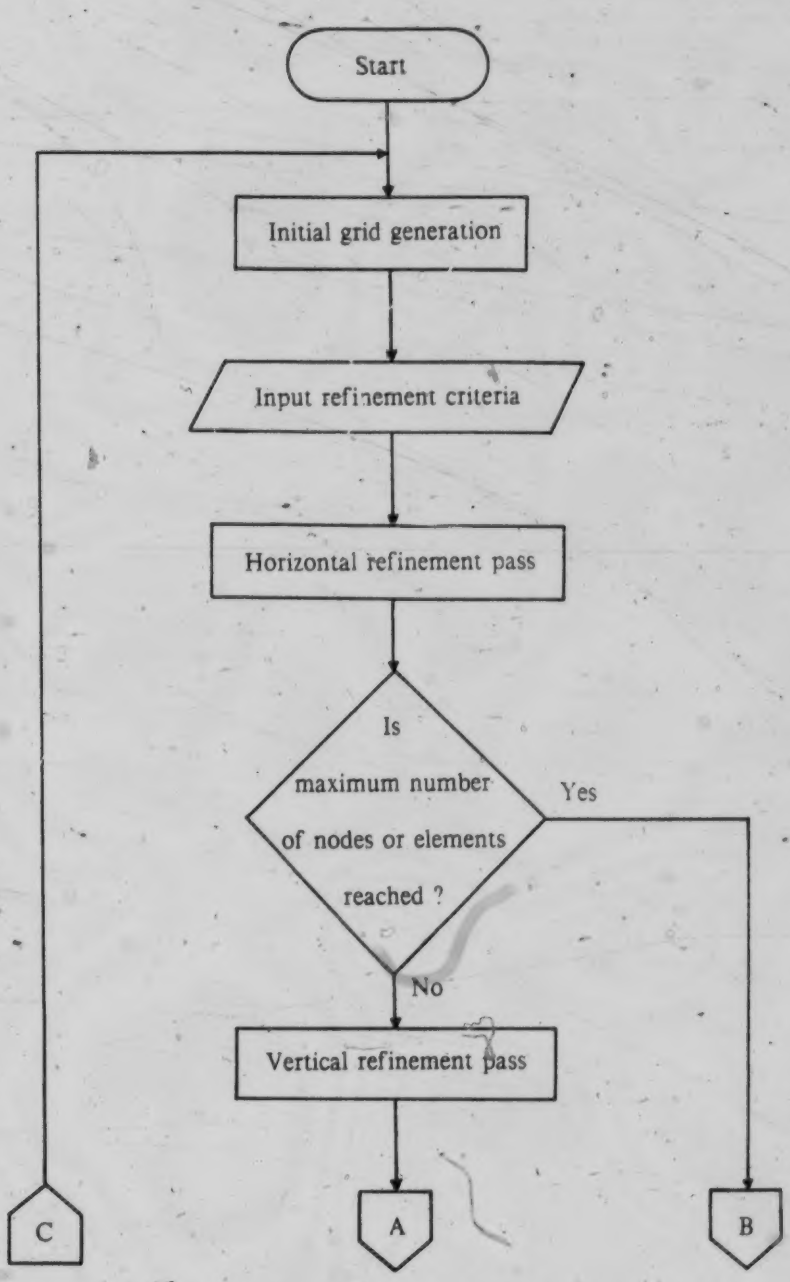


Fig. 4.7 (a) Flowchart for grid generation

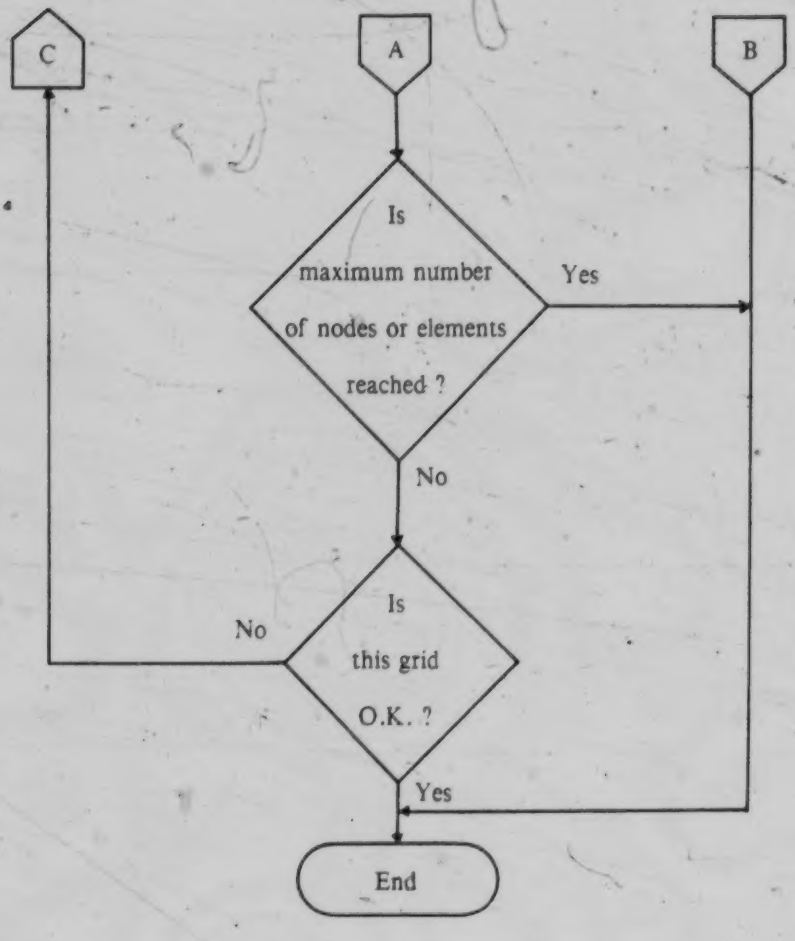


Fig. 4.7 (b) Flowchart for grid generation

Fig. 5.1 Element types in one dimension [37]

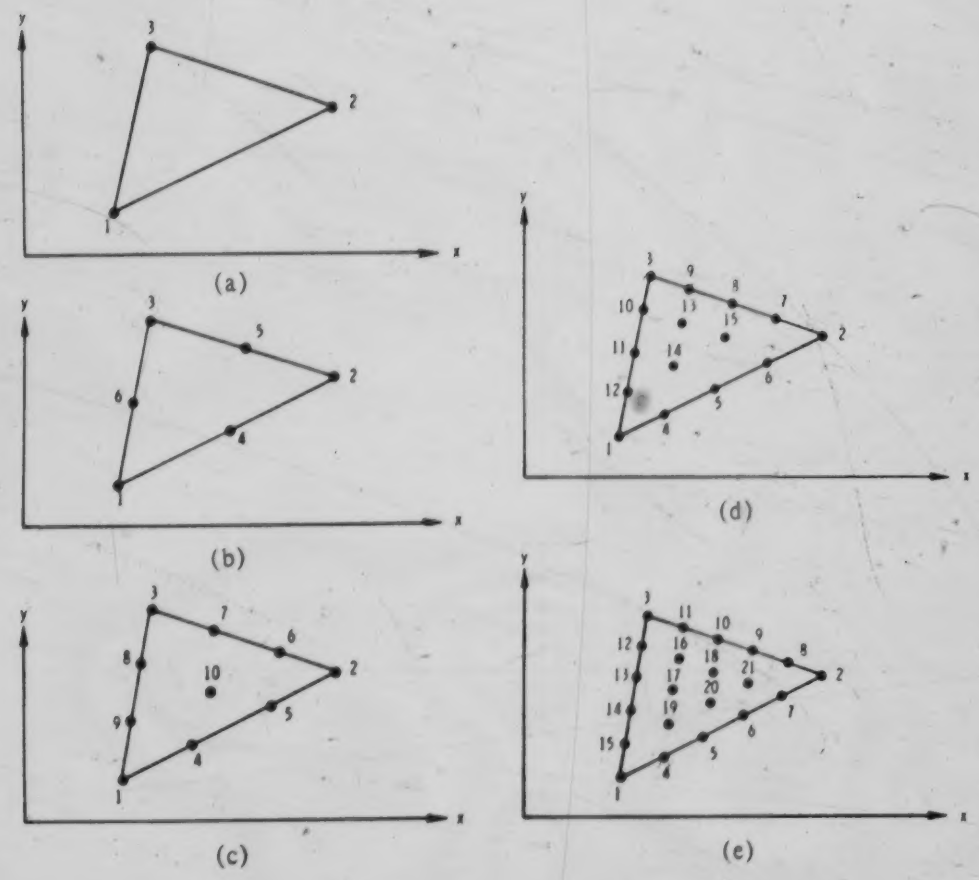


Fig. 5.2 Element types in two dimension with ϕ specified at the nodes [37]

(a)Linear (b)Quadratic (c)Cubic (d)Quartic (e)Quintic

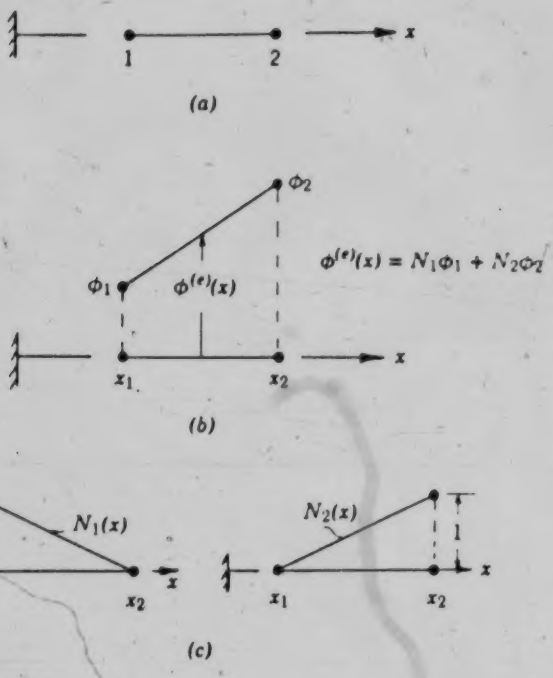


Fig. 5.3 Linear interpolation functions for two node line elements [37]

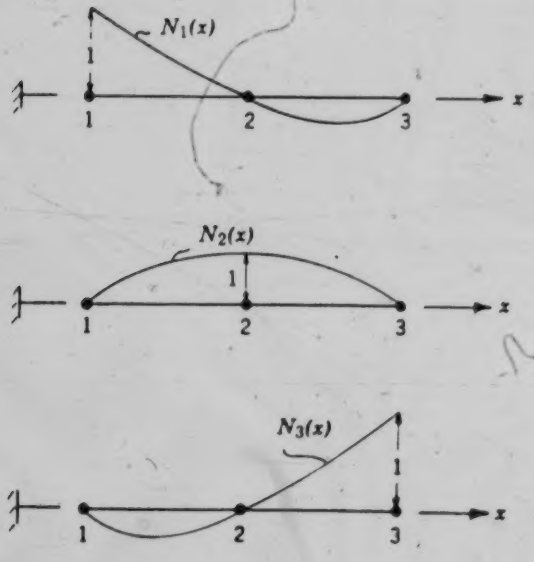


Fig. 5.4 Quadratic interpolation functions for three node line elements [37]

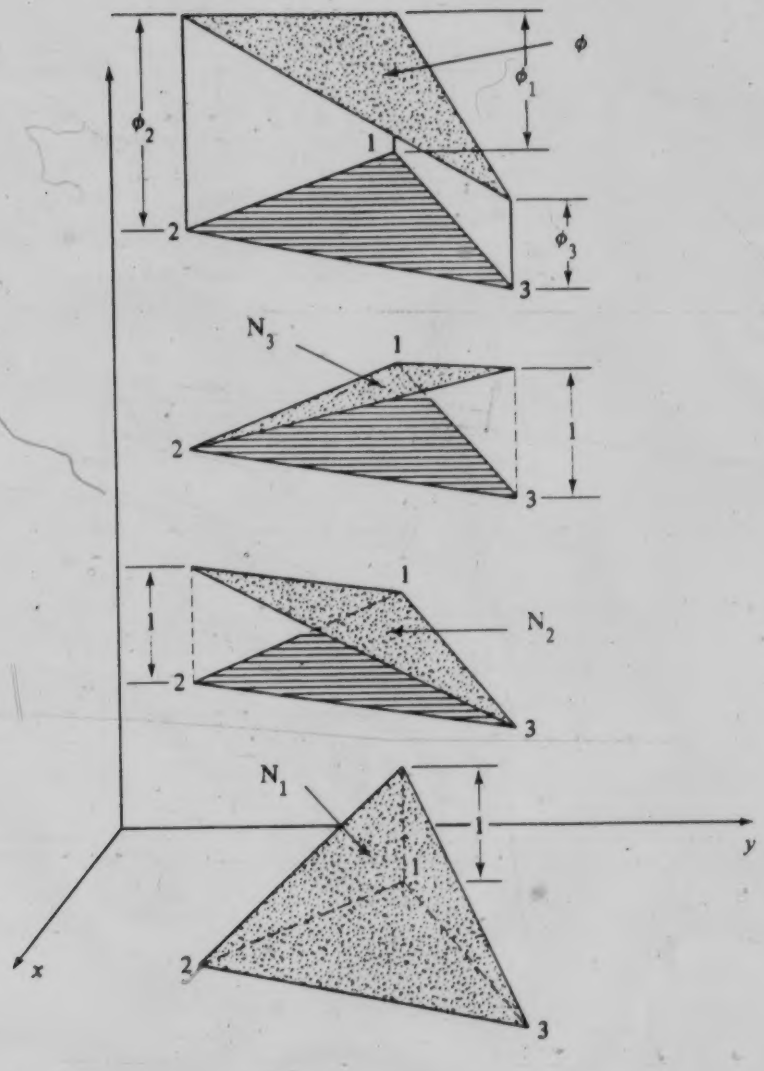
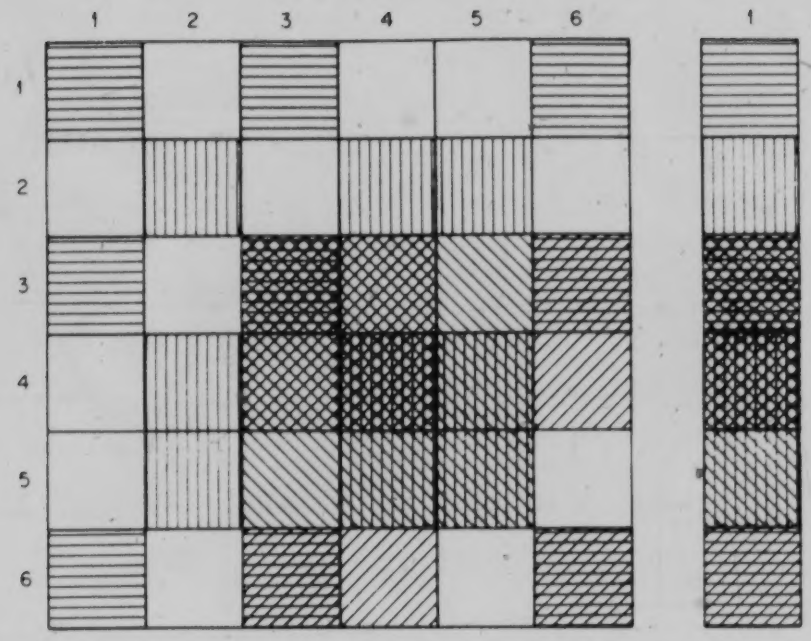
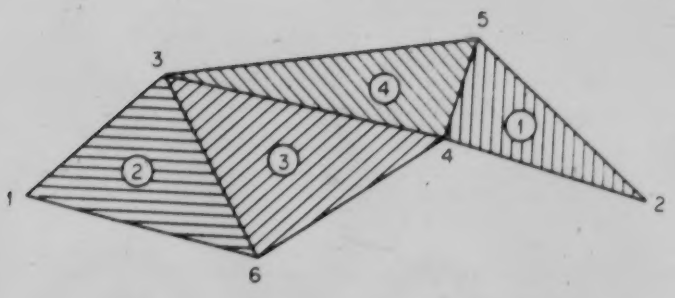


Fig. 5.5 Linear interpolation functions for a three node triangular element [46]

$$\phi = N_1\phi_1 + N_2\phi_2 + N_3\phi_3$$



(a)



(b)

Fig. 5.6 Assembled contributions $K\phi = Q$ [47]

(a) System matrices (b)Contributing elements

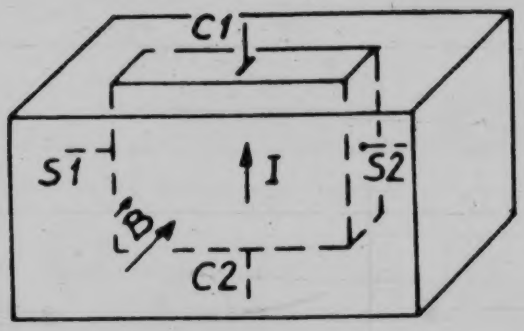


Fig. 6.1 (a) The conventional Hall plates [48]

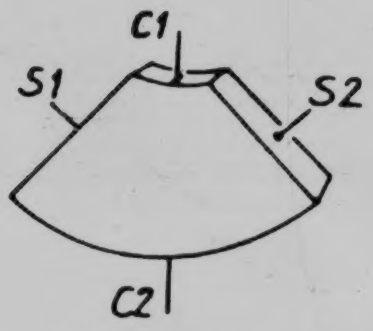


Fig. 6.1 (b) Partly deformed Hall plates [48]

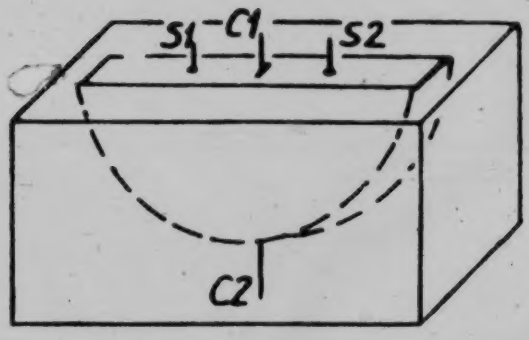


Fig. 6.1 (c) Semicircular Hall plates [48]

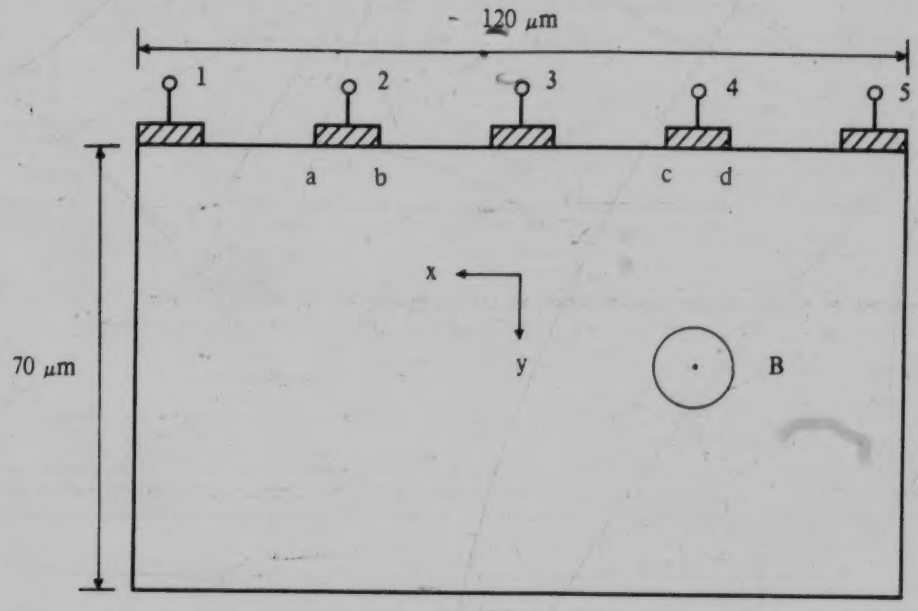


Fig. 6.2 Device geometry for the simulation of the VHD

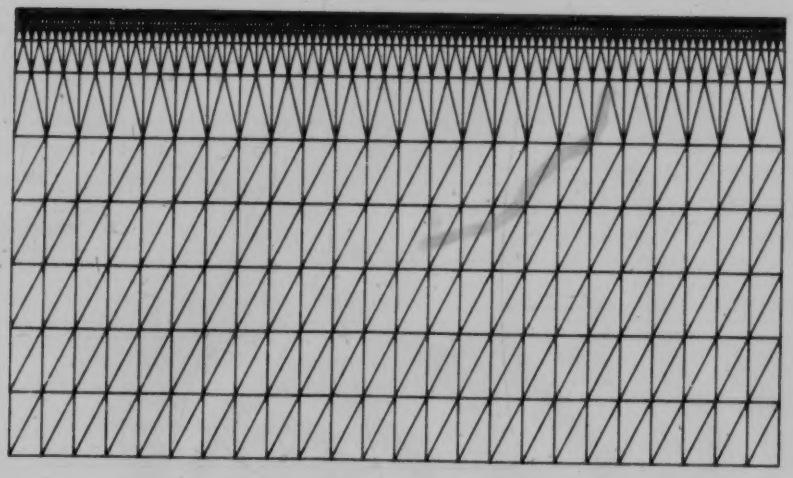
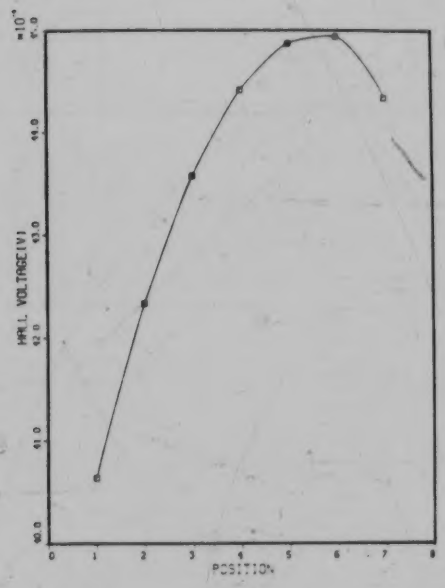
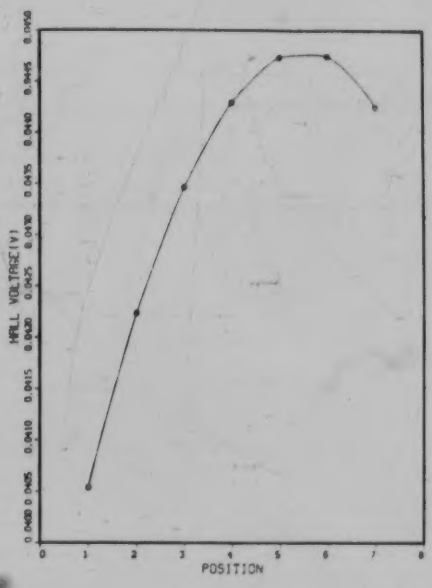


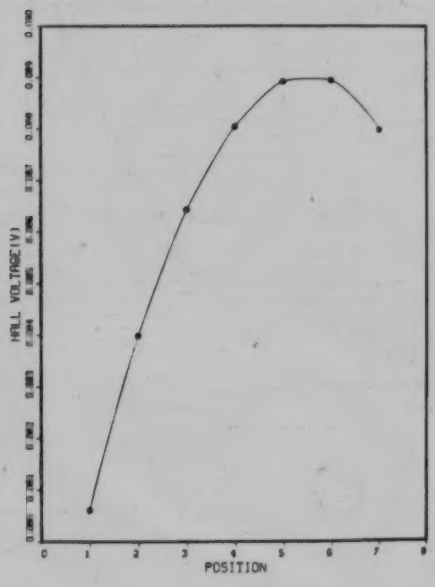
Fig. 6.3 Grid for the simulation of the VHD



(a)



(b)



(c)

Fig. 6.4 The Hall voltage vs. sensor probe positions

(a) μ*B = 0.01

(b) μ*B = 0.1

(c) μ*B = 0.2

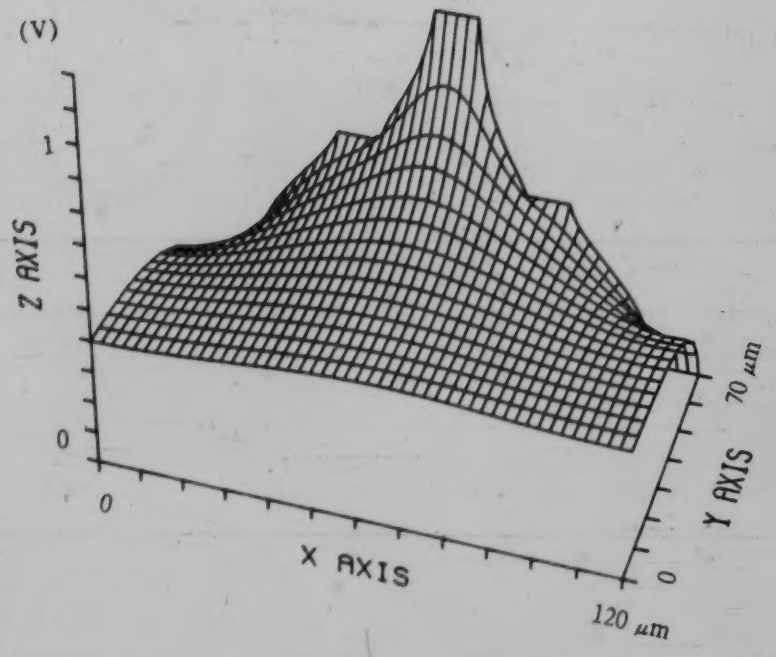


Fig. 6.5 Potential distribution at optimum position (μ*B = 0.2)

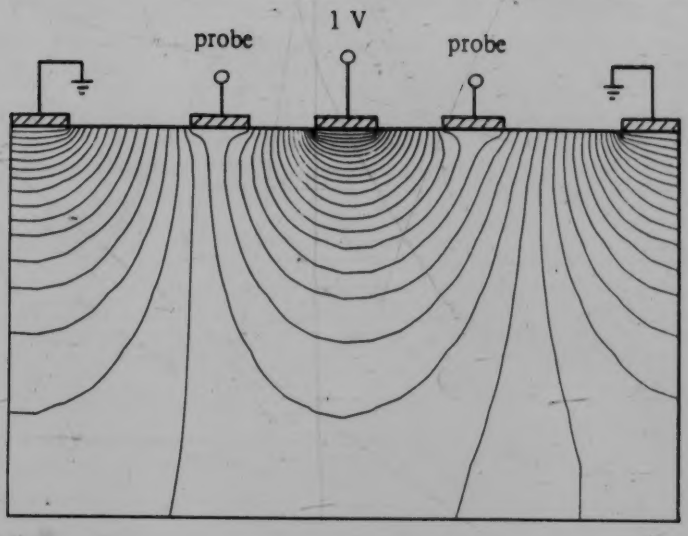


Fig. 6.6 Equipotential lines at optimum position (μ*B = 0.2)

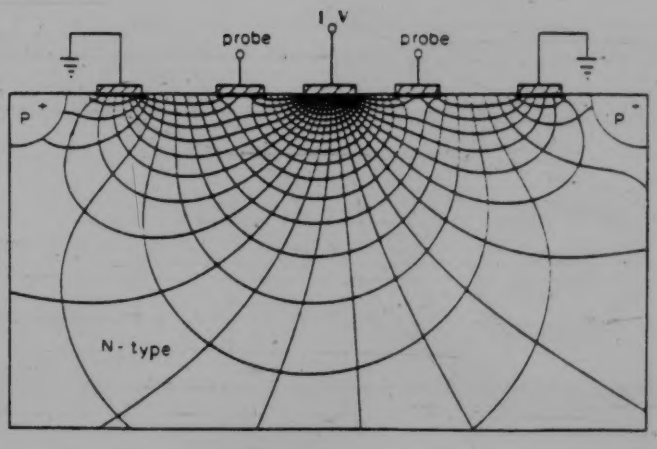


Fig. 6.7 Equipotential and current lines from Huiser and Baltes [35,36]

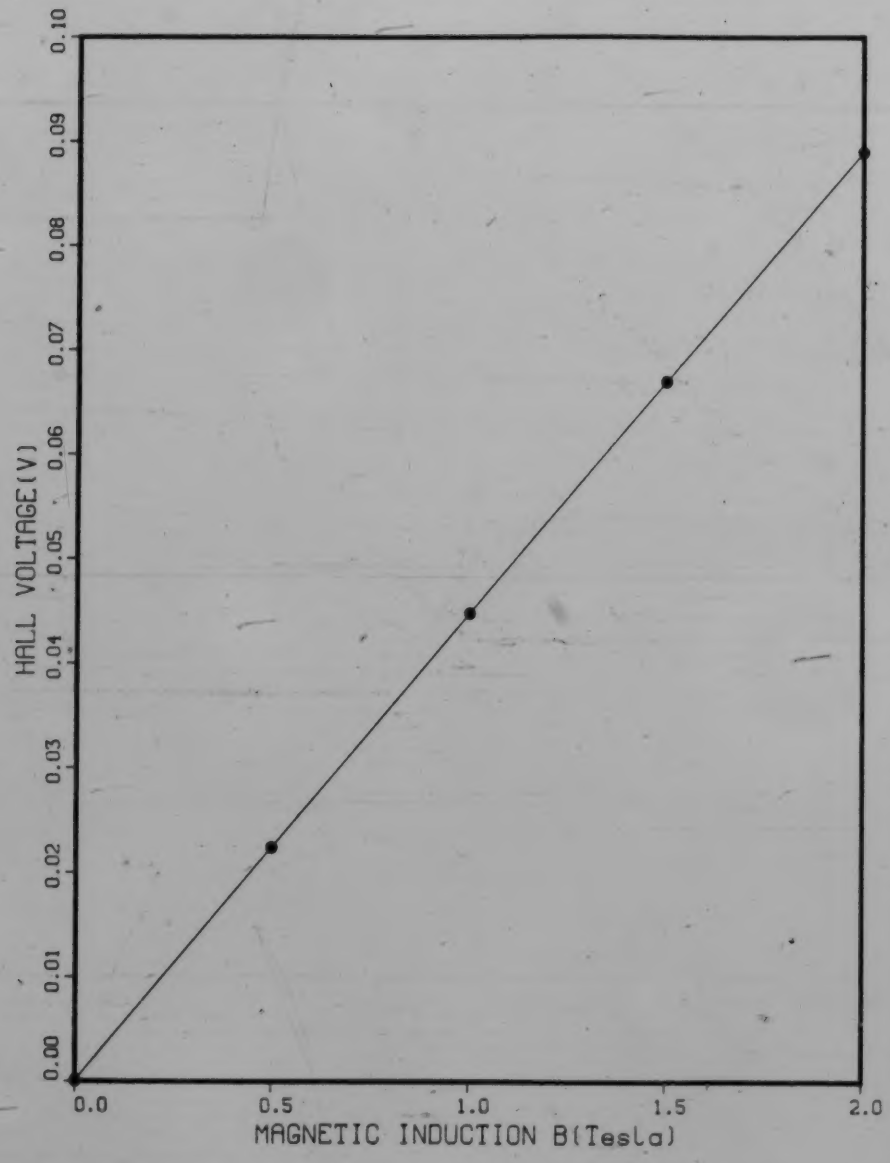


Fig. 6.8 The Hall voltage vs. magnetic induction

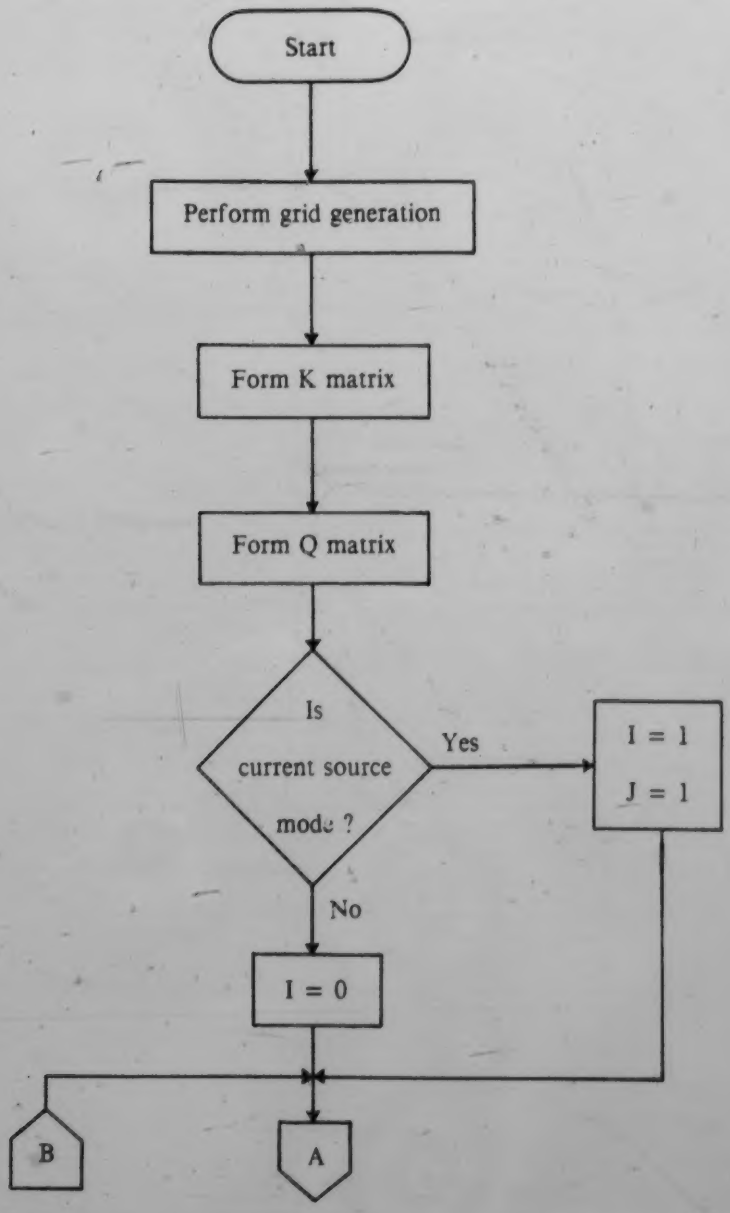


Fig. 6.9 (a) Flowchart for the simulation of the VHD

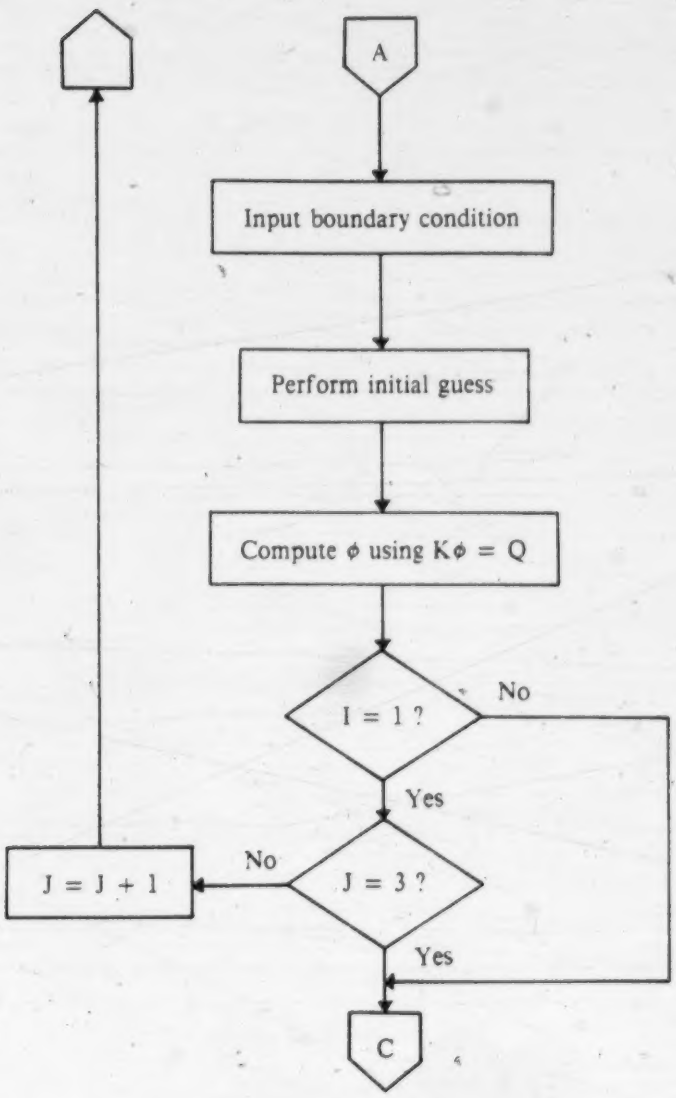


Fig. 6.9 (b) Flowchart for the simulation of the VHD

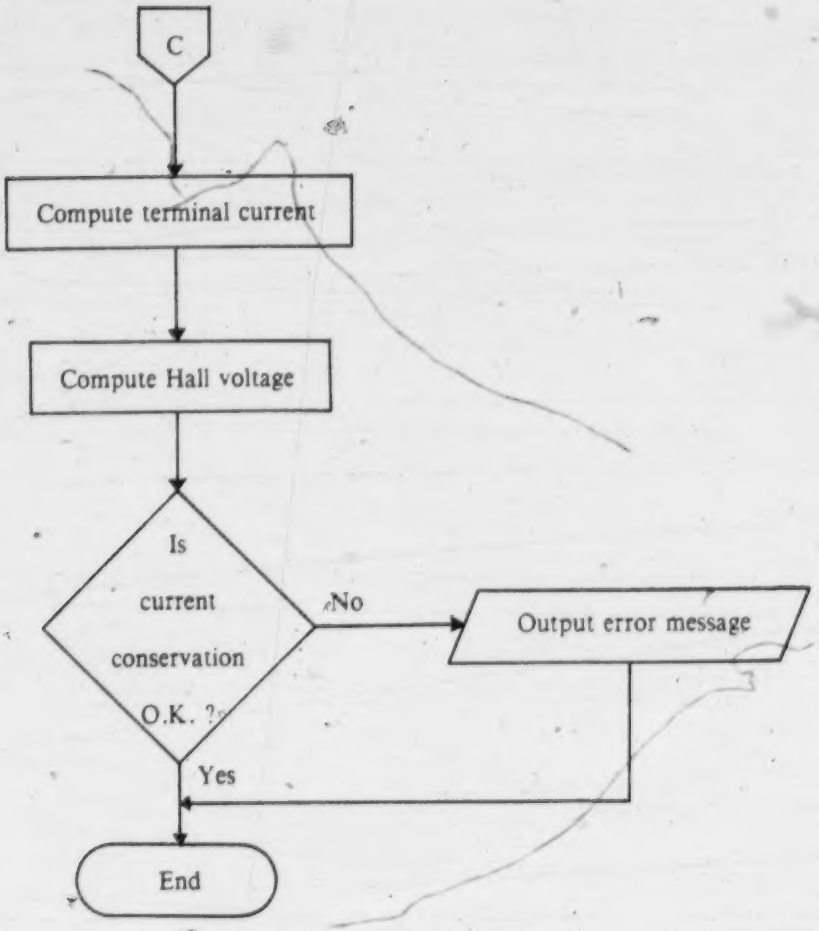


Fig. 6.9 (c) Flowchart for the simulation of the VHD

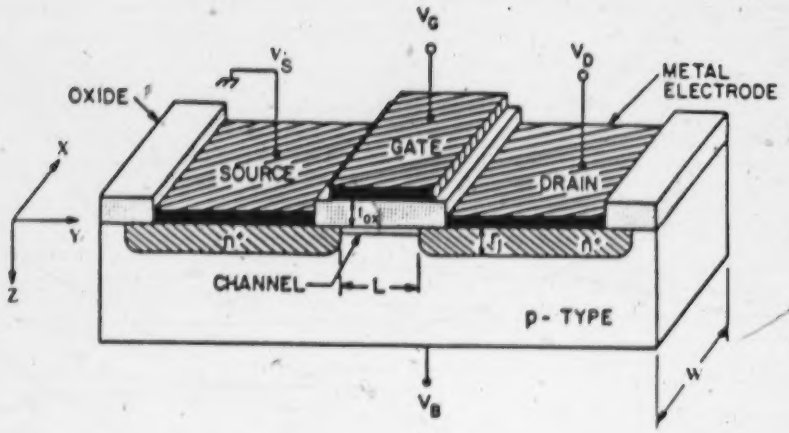


Fig. 7.1 Schematic of an n-channel MAGFET [16]

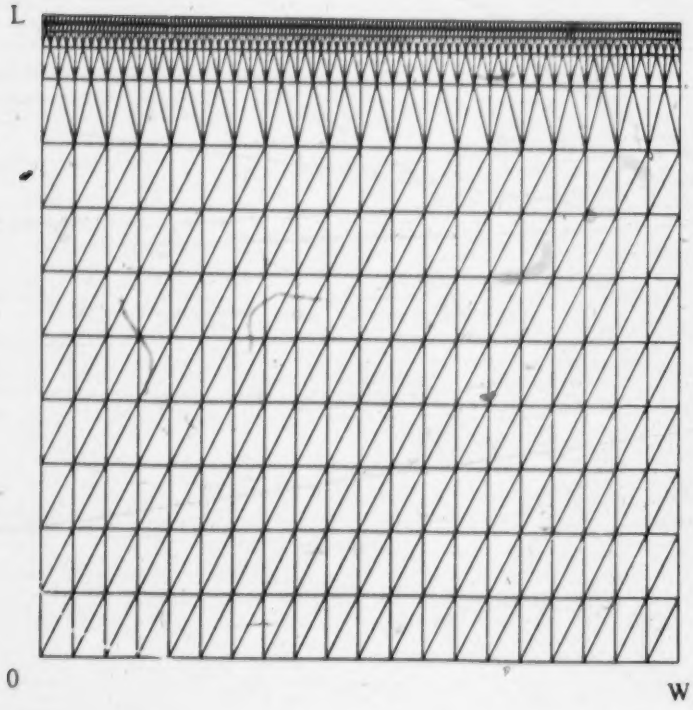


Fig. 7.2 Grid for the simulation of the MAGFET (W = L = 100 μm)

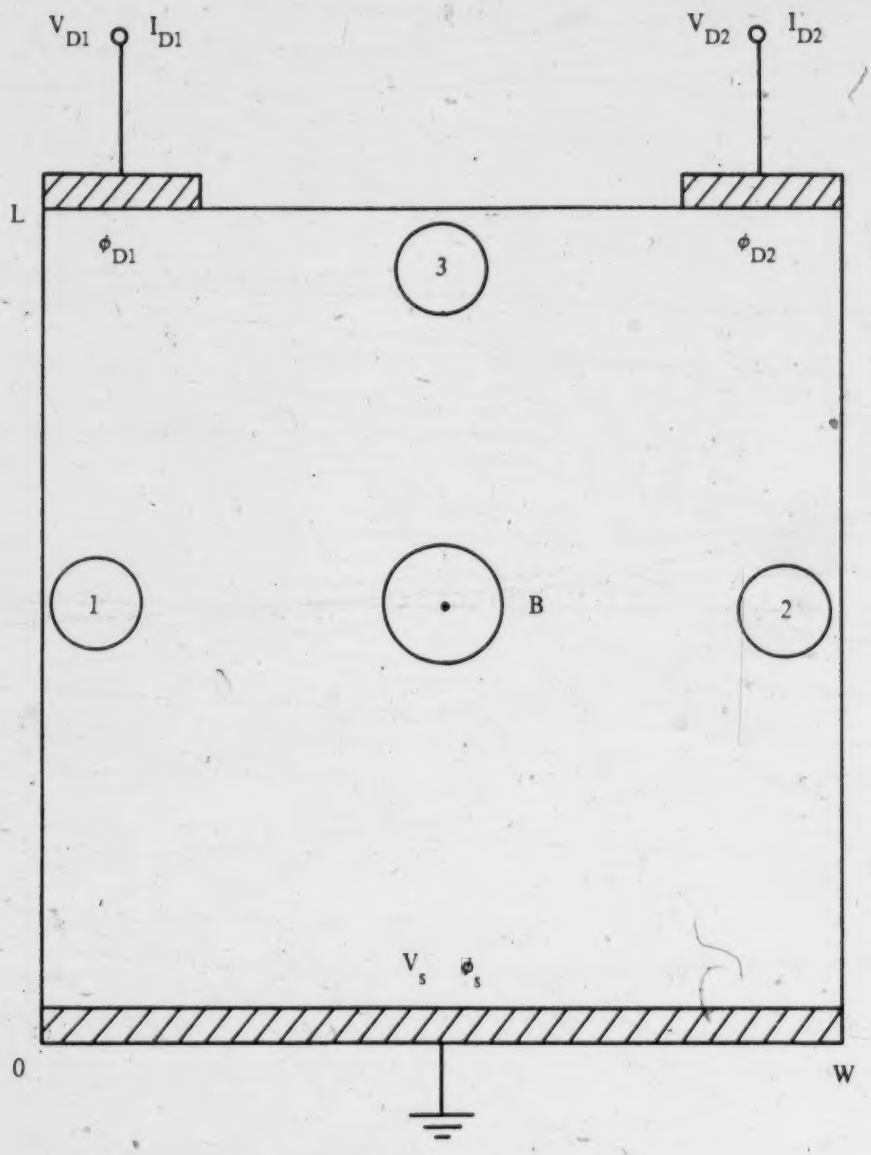


Fig. 7.3 Top view of the double drain MAGFET

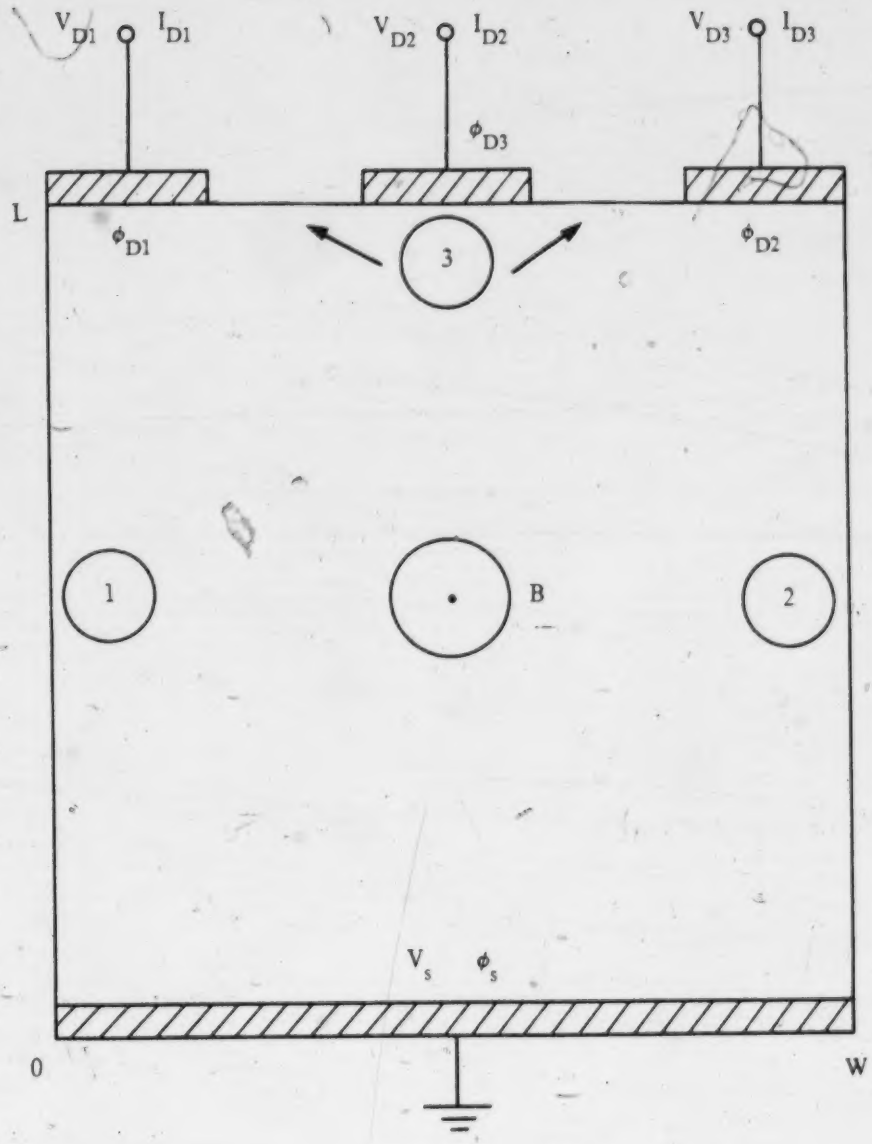


Fig. 7.4 Top view of the triple drain MAGFET

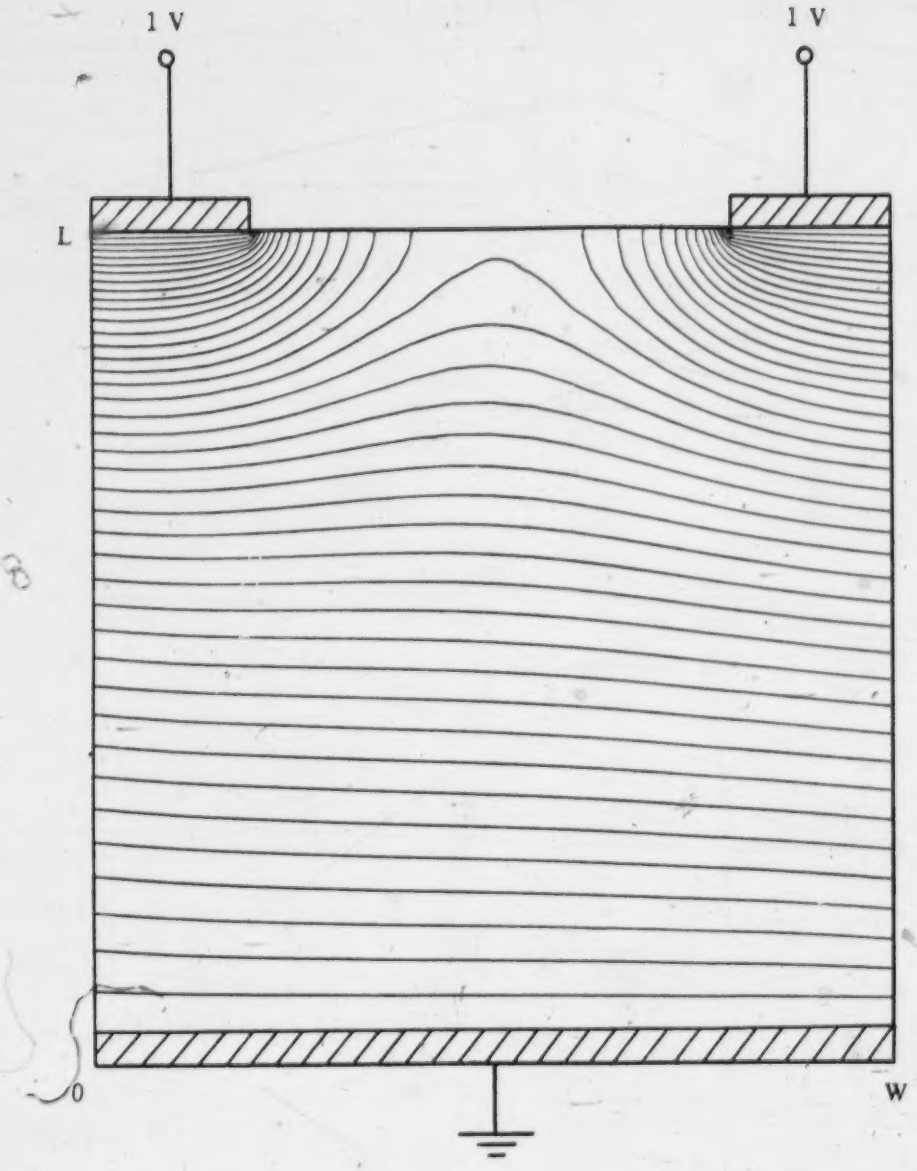


Fig. 7.5 Equipotential lines of the double drain MAGFET

(μ*B = 0.1, W = L = 100 μm)

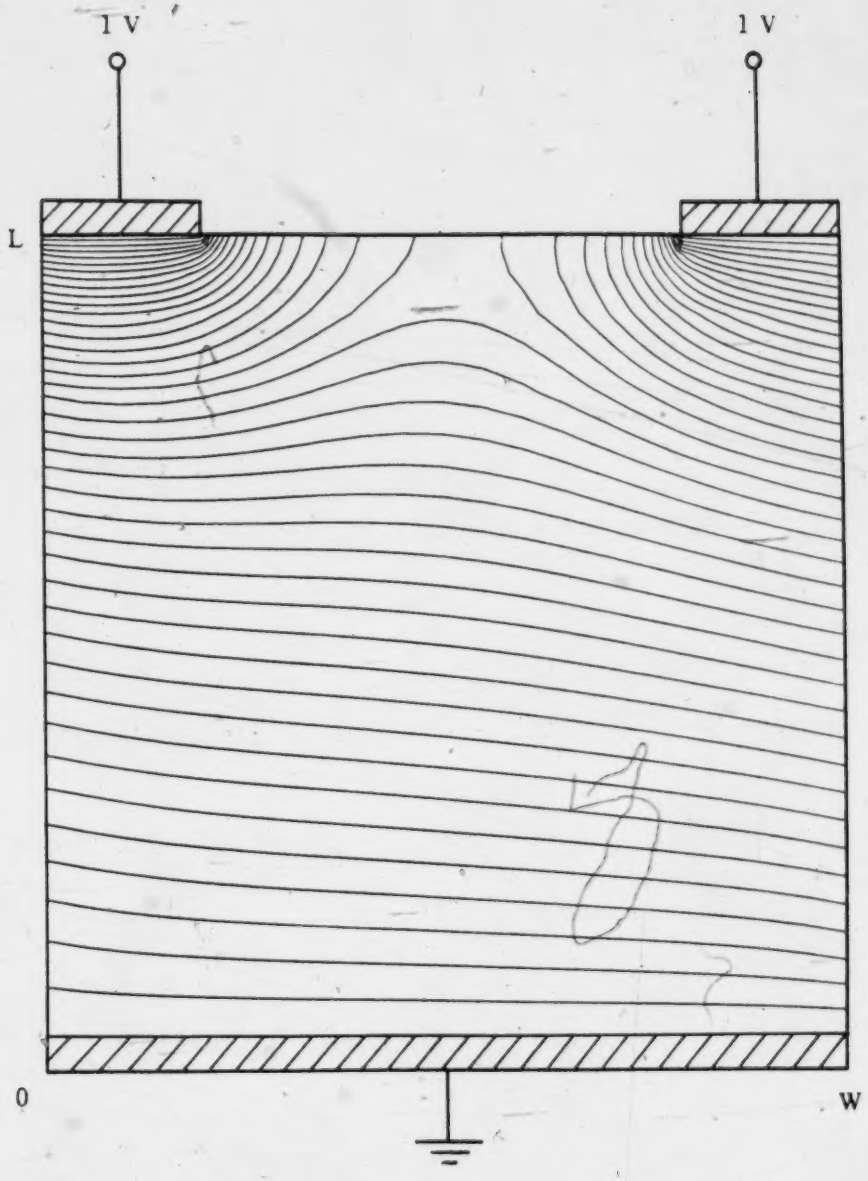


Fig. 7.6 Equipotential lines of the double drain MAGFET

(μ*B = 0.2, W = L = 100 μm)

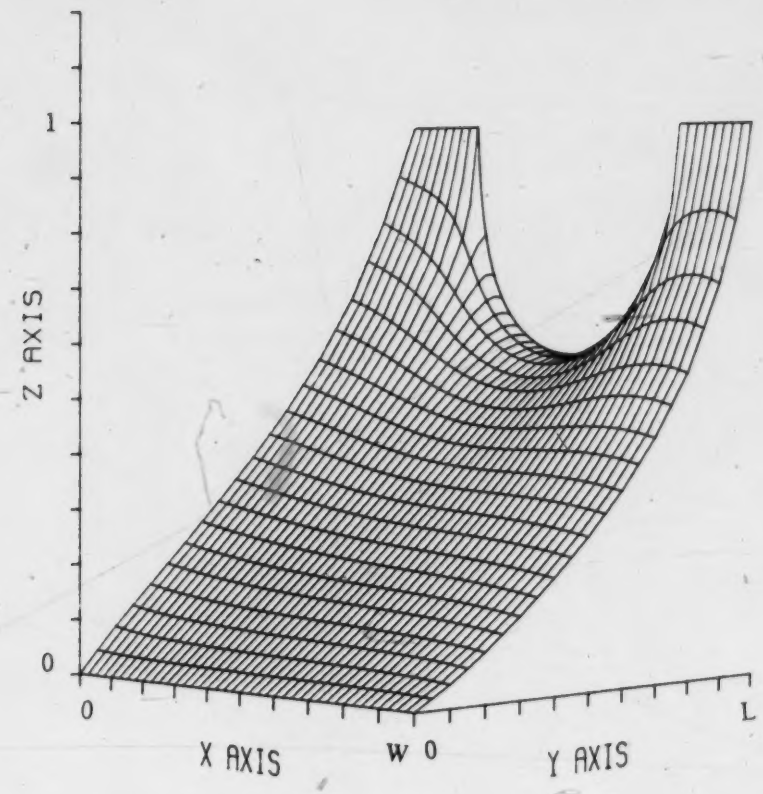


Fig. 7.7 Potential distribution of the double drain MAGFET

(μ*B = 0.2, W = L = 100 μm)

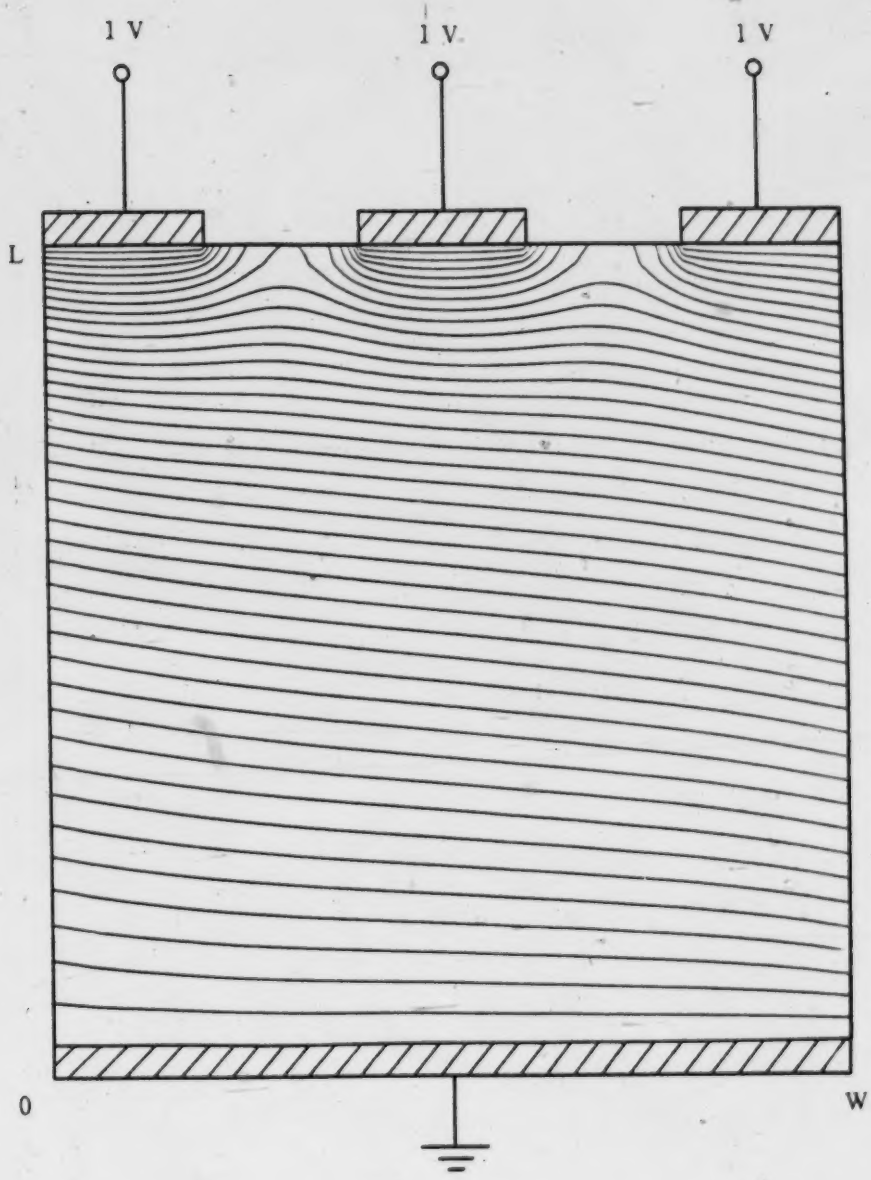


Fig. 7.8 Equipotential lines of the triple drain MAGFET

(μ*B = 0.2, W = L = 100 μm)

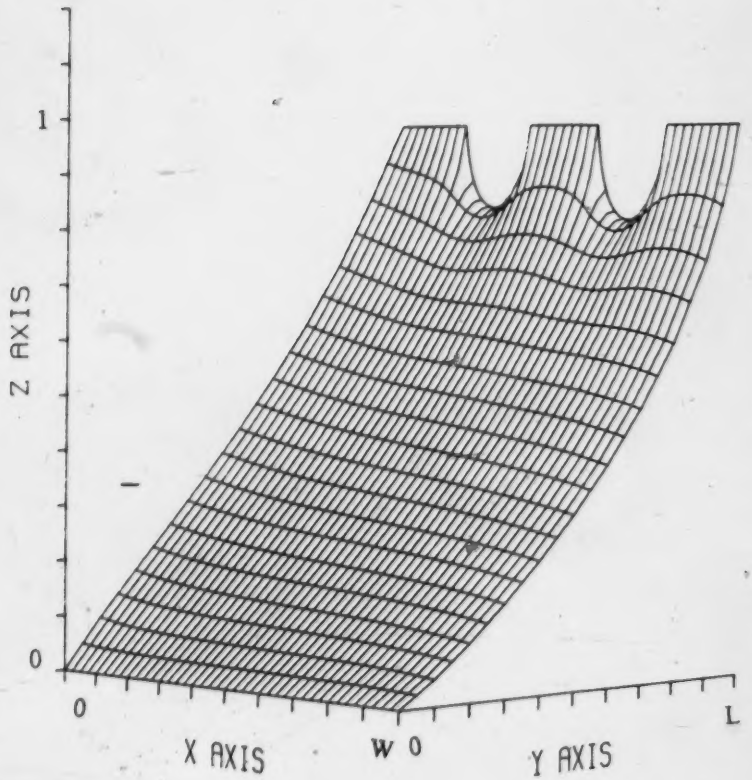


Fig. 7.9 Potential distributions of the triple drain MAGFET

(μ*B = 0.2, W = L = 100 μm)

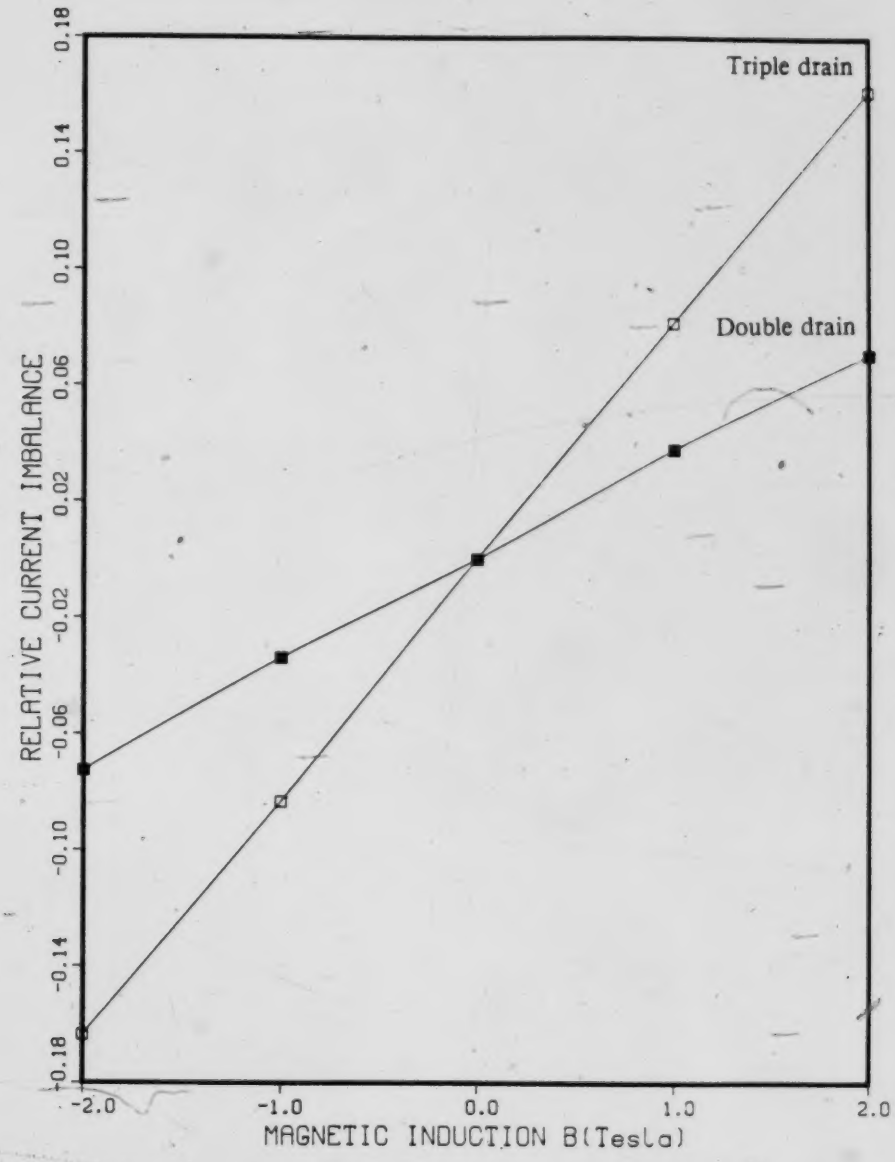


Fig. 7.10 Drain current imbalance vs. magnetic induction

(W = L = 100 μm)

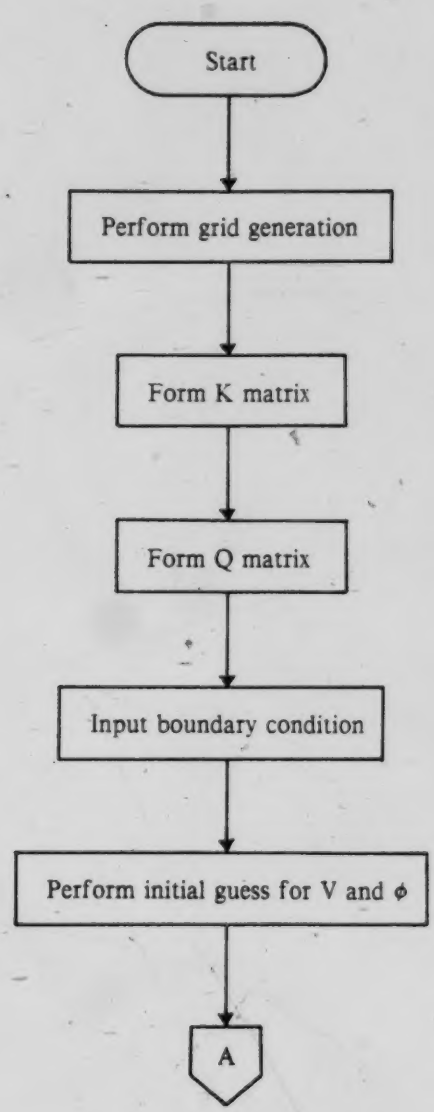


Fig. 7.11 (a) Flowchart for the simulation of the MAGFET

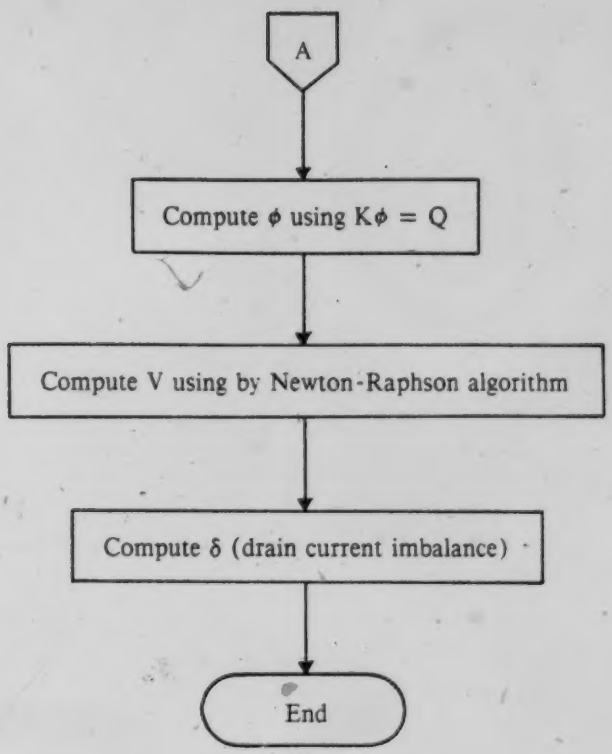


Fig. 7.11 (b) Flowchart for the simulation of the MAGFET

9. REFERENCES

[1] H. P. Baltes, 'Magnetic Sensors'(invited), *European Physics Society, 6th General Conference "Trends in Physics"*, Prague, Aug. 27-31, 1984, Technical Digest, p. 513 and Proc. pp. 606-611.

[2] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*, New York, Springer-Verlag, 1984.

[3] H. P. Baltes and R. S. Popović, 'Integrated semiconductor magnetic field sensors', *Proc. IEEE*, in press.

[4] V. Zieren, 'Integrated silicon multicollector magnetotransistors', *Ph.D Thesis*, Delft Technical University, 1983.

[5] A. Nathan, 'Numerical analysis of MOS Magnetic Field Sensors', *M.Sc Thesis*, The University of Alberta, 1984.

[6] R. C. Gallagher and W. S. Corak, 'A metal-oxide-semiconductor (MOS) Hall element', *Solid-State Electronics*, vol. 9, pp. 571-580, 1966.

[7] M. Hirata and S. Suzuki, 'Integrated Magnetic Sensor', *Proc. 1st Sensor Symposium*, pp. 305-310, 1981.

[8] P. W. Fry and S. J. Hoey, 'A silicon MOS magnetic field transducer of high sensitivity', *IEEE Trans. Electron Devices*, vol. ED-16, no. 1, pp. 35-39, 1969.

[9] R. S. Popovic and H. P. Baltes, 'A CMOS Magnetic Field Sensor', *IEEE J. Solid-State Circuits*, vol. SC-18, pp. 426-428, 1983.

[10] V. Zieren and B. P. M. Duyndam, 'Magnetic-Field-Sensitive Multicollector n-p-n Transistors', *IEEE Trans. Electron Devices*, vol. ED-29, no. 1, pp. 83-90, 1982.

[11] V. Zieren, S. Kordíc and S. Middelhoek, 'Comment on "Magnetic Transistor Behavior Explained by Modulation of Emitter Injection, not Carrier Deflection"', *IEEE Electron Device Lett.*, vol. EDL-3, no. 12, pp. 394-395, 1982.

[12] H. Pfleiderer, 'Magnetodiode model', *Solid State Electronics*, vol. 15, pp. 335-353, 1972.

[13] G. Kamarinos, P. Viktorovitch, S. Christoloveneau, J. Borrel and R.Staderini, 'Silicon on sapphire magnetodiodes of high sensitivities', *IEDM Dig. Tech. papers*, Washington, D.C.(USA), pp. 114a-114c, 1977.

[14] C. Kittel, *Introduction to Solid State Physics*, 2nd ed., New York, Wiley, 1962.

[15] K. Seeger, *Semiconductor Physics : An introduction*, 2nd ed., New York, Spring-Verlag, 1982.

[16] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed., New York, Wiley, 1981.

[17] O. Madelung, *Introduction to Solid-State Theory*, Berlin, Springer-Verlag, 1978.

[18] J. G. Shaw, 'Numerical Analysis of Semiconductor Devices', *Ph.D Thesis*, The University of Manitoba, 1983.

[19] E. Arnold, 'Computer simulation of conductivity and Hall effect in inhomogeneous inversion layers', *Surface science*, vol. 113, pp. 239-243, 1982.

[20] M. Kurata, *Numerical Analysis for Semiconductor Devices*, Lexington, MA:Heath, 1982.

[21] W. L. Engl, H. K. Dirks and B. Meinerzhagen, 'Device Modeling', *Proc. IEEE*, vol. 71, no. 1, pp. 10-33, 1983.

[22] W. Fichtner, R. E. Bank and D. J. Rose, 'Semiconductor Device Simulation', *IEEE Trans. Electron Devices*, vol. ED-30, no. 9, pp. 1018-1030, 1983.

[23] R. E. Bank, D. J. Rose and W. Fichtner, 'Numerial Methods for Semiconductor Devie Simulation', *IEEE Trans. Electron Devices*, vol. ED-30, no. 9, pp. 1031-1041, 1983.

[24] Proc. NASECODE III, Galway, *Numerical Analysis of Semiconductor Devices and Integrated Circuits*, ed. by J. J. H. Miller, Boole Press, Dublin, 1983.

[25] H. P. Baltes, L. Andor, A. Nathan and H. G. Schmidt-Weinmar, 'Two-dimensional numerical analysis of a silicon magnetic field sensor', *IEEE Trans. Electron Devices*, vol. ED-31, no. 7, pp. 996-999, 1984.

[26] L. Andor, H. P. Baltes, A. Nathan and H. G. Schmit-Weinmar, 'Carrier transport in semiconductor magnetic field sensors', *IEDM 1983, Technical Digest*, pp. 635-638.

[27] H. G. Schmidt-Weinmar, L. Andor, H. P. Baltes and A. Nathan, 'Numerical modeling of silicon magnetic field sensors : magnetoconcentration effects in split-metal-contact devices', *IEEE Trans. Magnetics*, vol. MAG-20, no. 5, pp. 975-977, 1984.

[28] H. P. Baltes, L. Andor, A. Nathan and H. G. Schmidt-Weinmar, 'Advances in modeling magnetic semiconductor devices', *Proc. MIEL-84*, pp. 104-116, May 7-9, Nis, Yugoslavia, 1984.

[29] A. Nathan, L. Andor, H. P. Baltes and H. G. Schmidt-Weinmar, 'Modeling of a Dual-Drain NMOS Magnetic-Field Sensor', *IEEE J. Solid-State Circuits*, vol. SC-20, no. 3, pp. 819-821, 1985.

[30] L. Andor, H. P. Baltes, A. Nathan and H. G. Schmidt-Weinmar, 'Numerical modeling of magnetic-field-sensitive semiconductor devices', *IEEE Trans. Electron Devices*, vol. ED-32, no. 7, pp. 1224-1230, 1985.

[31] A. Nathan, A. M. J. Huiser, H. P. Baltes and H. G. Schmidt-Weinmar, 'A triple drain MOSFET magnetic field sensor', *Second Canadian Semiconductor Technology Conference*, Ottawa, Aug. 22-24, pp. 27-28, 1984.

[32] A. Nathan, A. M. J. Huiser and H. P. Baltes, 'Two-dimensional numerical modeling of magnetic-field sensors in CMOS technology', *IEEE Trans. Electron Devices*, vol. ED-32, no. 7, pp. 1212-1218, 1985.

[33] D. L. Scharfetter and H. K. Gummel, 'Large-signal analysis of a silicon read diode oscillator', *IEEE Trans. Electron Devices*, vol. ED-16, no. 1, pp. 64-77, 1969.

[34] Y. Sugiyama, 'Fundamental research on Hall effects in inhomogeneous magnetic fields', *Researches of the Electrotechnical Laboratory*, No. 838, Tokyo: Electrotechnical Laboratory, 1983.

[35] A. M. J. Huiser and H. P. Baltes, 'Numerical Modeling of Vertical Hall-Effect Devices', *IEEE Electron Device Lett.*, vol. EDL-5, no. 11, pp. 482-484, 1984.

[36] A. M. J. Huiser and H. P. Baltes, 'A 3-D Numerical Simulation of a Vector Hall Device', *Proc. 1984 Canadian Conference on Very Large Integration*, pp. 4.110-4.113, 1984.

[37] K. H. Huebner and E. A.Thornton, *The Finite Element Method for Engineers*, 2nd ed., New York, Wiley, 1982.

[38] J. J. Barnes and R. J. Lomax, 'Finite-Element Methods in Semiconductor Device Simulation', *IEEE Trans. Electron Devices*, vol. ED-24, no. 8, pp. 1082-1089, 1977.

[39] R. Guerrieri and M. Rudan, 'Optimization of the Finite-Element Solution of the Semiconductor-Device Poisson Equation', *IEEE Trans. Electron Devices*, vol. ED-30, no. 9, pp. 1097-1103, 1983.

[40] G. Strang and G. J. Fix, *An Analysis of The Finite Element Method*, New Jersey, Prentice-Hall, 1973.

[41] J. R. Westlake, *A Handbook of Numerical Matrix Inversion and Solution of Linear Equations*, New York, Wiley, 1968.

[42] E. B. Becker, G. F. Carey and J. T. Oden, *Finite Element : Computational Aspects*, New Jersey, Prentice-Hall, 1984.

[43] M. Rudan and R. Guerrieri, 'Relevant problems in the numerical simulation of semiconductor devices using the finite element method', in : *Numerical Analysis of Semiconductor Devices and Integrated Circuits*, ed. by J. J. H. Miller, Boole Press, Dublin, 1983.

[44] E. B. Becker, G. F. Carey and J. T. Oden, *Finite Element : A Second Course*, New Jersey, Prentice-Hall, 1983.

[45] A. R. Mitchell and R. Wait, *The Finite Element Method in Partial Differential Equations*, London, Wiley, 1977.

[46] E. B. Becker, G. F. Carey and J. T. Oden, *Finite Element : An Introduction*, New Jersey, Prentice-Hall, 1981.

[47] T. E. Akin, *Application and Implementation of Finite Element Methods*, London, Academic Press, 1982.

[48] R. S. Popović, 'The Vertical Hall-Effect Device', *IEEE Electron Devices Lett.*, vol. EDL-5, no. 9, pp. 357-358, 1984.

[49] W. Allegretto, Y. S. Mun, A. Nathan and H. P. Baltes, 'Optimization of Semiconductor Magnetic Field Sensors Using Finite Element Analysis', *Proc.*

NASECODE IV Conf., Dublin: Boole Press 1985, in press.

[50] A. Nathan, Y. S. Mun, W. Allegretto and H. P. Baltes, 'Numerical Simulation of Semiconductor Magnetic Field Sensors Using Finite-Element Method', *Applied Simulation and Modeling ASM '85*, Proc. ed. by D. Koval and M. H. Hamza, Acta Press, Annaheim, pp. 22-24, 1985.

[51] R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*, New York, Wiley, 1977.

[52] S. D. Conte and C. de Boor, *Elementary Numerical Analysis; an algorithmic approach*, 3rd ed., New York, McGraw-Hill, 1980.

[53] M. Rudan, 'Accurate Charge-Integrals Evalution in Two-Dimensional Finite-Element Semiconductor Device Simulation', *Proc. COMPEL*, vol. 1, no. 1, pp. 37-51, 1982.

10. APPENDIX

If the coordinates of three corner points of a triangle are (x_1,y_1), (x_2,y_2) and (x_3,y_3), then the area of this triangle Δ is [37,53];

$$\Delta = \frac{1}{2}\begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix} . \qquad (A.1)$$

The integration of natural coordinates,often called area coordinates, which are precisely the interpolation functions for linear interpolation over a triangle is the following [37];

$$\int_D N_1^{\alpha} N_2^{\beta} N_3^{\omega} \, dD = \frac{\alpha!\ \beta!\ \omega!}{(\alpha + \beta + \omega + 2)!}\ 2\Delta . \qquad (A.2)$$